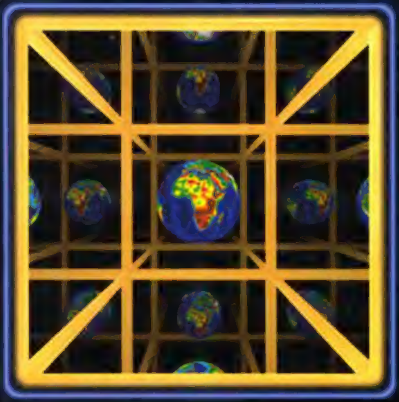
The SHAPE of SPACE

Second Edition



JEFFREY R. WEEKS

Recipient of the MacArthur Fellowship

The SHAPE of SPACE

Second Edition

Jeffrey R. Weeks
Freelance Mathematician
Canton, New York

MARCEL DEKKER

MARCEL DEKKER, INC. NEW YORK • BASEL

ISBN: 0-8247-0709-5

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-261-8482; fax: 41-61-261-8896

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

To Nadia

This Page Intentionally Left Blank

Preface to the Second Edition

When the first edition of *The Shape of Space* appeared in 1985, the idea of measuring the shape of the real universe was only a pleasant dream that I hoped might be realized in the distant future. That future has arrived, and it came sooner than expected. As of 2002, two independent research projects are underway that attempt to measure the shape of space in different ways. The method of Cosmic Crystallography (Chapter 21) looks for patterns in the arrangement of the galaxies, while the Circles in the Sky method (Chapter 22) uses microwave radiation remaining

from the big bang. It's too soon to say whether either method will succeed, but there is no doubt that the first decade of the 21st century marks humanity's first viable attempt to measure the shape of space.

In contrast to our rapidly evolving knowledge of the physical universe, our understanding of the basic geometry of surfaces and three-dimensional manifolds was already mature in 1985 and has changed little since then. Therefore, the first 18 chapters of this book follow the same line of development as in the first edition. The biggest improvement is to fill a logical gap. When writing the first edition of this book I couldn't find a sufficiently simple proof of the classification of surfaces. During the intervening years John Conway devised his ingenious ZIP proof. Appendix C reproduces an elementary exposition of the ZIP proof featuring George Francis's superbly clear illustrations.

Finally, the bibliography in Appendix B has been brought up to date.

I wish you well in your exploration of strange spaces, and I hope you have as much fun with them as I have.

Jeffrey R. Weeks

Preface to the First Edition

Möbius strips and Klein bottles first caught my interest when I was in high school. I knew they were part of something called "topology" and I was eager to learn more. Sadly, neither the school library nor the public one had much on the subject. Perhaps, I thought, I will learn more about topology in college. In college I couldn't even sign up for topology until my senior year, and even then all I got was one course in extreme generalization (point-set topology) and another that developed a collection of technical tools (algebraic to-

pology). Topology's most beautiful examples got bypassed completely.

The Shape of Space fills the gap between the simplest examples, such as the Möbius strip and the Klein bottle, and the sophisticated mathematics found in upper-level college courses. It is intended for a wide audience: I wrote it mainly for the interested non-mathematician (perhaps a high school student who has heard of Möbius strips and wants to learn more), but it also provides the intuitive examples that are currently missing from the college and graduate school curriculum.

I still haven't told you what the book is about, or even explained the title. For that I refer you to Chapter 1.

Jeffrey R. Weeks

Acknowledgments

This book never would have gotten anywhere without the help of the people who read the various drafts and suggested improvements. Particularly useful were the comments of Beth Christie, Bill Dunbar, Charles Harris, Geoff Klineberg, Skona Libowitz, Robert Lupton, Marie McAllister, Thea Pignataro, Evan Romer, Bruce Solomon, Harry Voorhees, Eric White, and an anonymous reviewer of the first draft. I would also like to thank the Princeton University math department for providing computer resources, George Francis for giving me drawing lessons (both in person and by mail), Nadia Marano and Craig Shaw for serving as art crit-

ics, Rudy Rucker for suggesting the title, Bill Thurston for inventing the mathematics described in Chapter 18, and my geometry students at Stockton State College for being the guinea pigs.

Natalia Panczyszyn Kozlowski painted Figure 7.4. The painting shows the view in a three-torus. Chapter 2 introduces the three-torus, and Chapter 7 explains the strange visual effects. The three-torus is a possible shape for the universe, one in which space is finite yet has no edges.

Figures 3.3, 9.8, 10.4, and 14.6 are from "The Mathematics of Three-dimensional Manifolds" by W. Thurston and J. Weeks (*Scientific American*, July 1984), copyright © 1985 by Scientific American, Inc., all rights reserved. Figures 4.9 and 5.6 are from Hilbert and Cohn-Vossen's *Anschauliche Geometrie* (English translation published by Chelsea, New York, 1952). Figure 13.8 is from Rudy Rucker's *Geometry, Relativity and the Fourth Dimension* (Dover, New York, 1977). Bill Thurston generously provided Figure 18.4. Richard Bassein prepared Figures 5.7, 5.8, 6.9, 6.10, 7.13, 8.2, 11.1, 11.4, 13.7, 15.1, 17.6, 18.1, and 18.2; John Heath did 2.4 through 2.8, and 4.8; Nadia Marano did 6.3; M. C. Escher did 8.1; Adam Weeks Marano did Figures 21.1, 21.2, 21.5, and 21.6; Figure 22.4 is ©Edward L. Wright, used with permission; and Figure 22.10 is the work of the MAP Science Team. George K. Francis drew the beautiful illustrations in Appendix C. The cover image is a screenshot from the interactive 3D software available for free at *www.northnet.org/weeks/SoS*.

Contents

The SHAPE *of* SPACE

This Page Intentionally Left Blank

The SHAPE *of* SPACE

PURE AND APPLIED MATHEMATICS

A Program of Monographs, Textbooks, and Lecture Notes

MONOGRAPHS AND TEXTBOOKS IN PURE AND APPLIED MATHEMATICS

1. *K. Yano,* Integral Formulas in Riemannian Geometry (1970)
2. *S. Kobayashi,* Hyperbolic Manifolds and Holomorphic Mappings (1970)
3. *V. S. Vladimirov,* Equations of Mathematical Physics (A. Jeffrey, ed.; A. Littlewood, trans.) (1970)
4. *B. N. Pshenichnyi,* Necessary Conditions for an Extremum (L. Neustadt, translation ed.; K. Makowski, trans.) (1971)
5. *L. Narici et al.,* Functional Analysis and Valuation Theory (1971)
6. *S. S. Passman,* Infinite Group Rings (1971)
7. *L. Dornhoff,* Group Representation Theory. Part A: Ordinary Representation Theory. Part B: Modular Representation Theory (1971, 1972)
8. *W. Boothby and G. L. Weiss, eds.,* Symmetric Spaces (1972)
9. *Y. Matsushima,* Differentiable Manifolds (E. T. Kobayashi, trans.) (1972)
10. *L. E. Ward, Jr.,* Topology (1972)
11. *A. Babakhanian,* Cohomological Methods in Group Theory (1972)
12. *R. Gilmer,* Multiplicative Ideal Theory (1972)
13. *J. Yeh,* Stochastic Processes and the Wiener Integral (1973)
14. *J. Barros-Neto,* Introduction to the Theory of Distributions (1973)
15. *R. Larsen,* Functional Analysis (1973)
16. *K. Yano and S. Ishihara,* Tangent and Cotangent Bundles (1973)
17. *C. Procesi,* Rings with Polynomial Identities (1973)
18. *R. Hermann,* Geometry, Physics, and Systems (1973)
19. *N. R. Wallach,* Harmonic Analysis on Homogeneous Spaces (1973)
20. *J. Dieudonné,* Introduction to the Theory of Formal Groups (1973)
21. *I. Vaisman,* Cohomology and Differential Forms (1973)
22. *B.-Y. Chen,* Geometry of Submanifolds (1973)
23. *M. Marcus,* Finite Dimensional Multilinear Algebra (in two parts) (1973, 1975)
24. *R. Larsen,* Banach Algebras (1973)
25. *R. O. Kujala and A. L. Vitter, eds.,* Value Distribution Theory: Part A; Part B: Deficit and Bezout Estimates by Wilhelm Stoll (1973)
26. *K. B. Stolarsky,* Algebraic Numbers and Diophantine Approximation (1974)
27. *A. R. Magid,* The Separable Galois Theory of Commutative Rings (1974)
28. *B. R. McDonald,* Finite Rings with Identity (1974)
29. *J. Satake,* Linear Algebra (S. Koh et al., trans.) (1975)
30. *J. S. Golan,* Localization of Noncommutative Rings (1975)
31. *G. Klambauer,* Mathematical Analysis (1975)
32. *M. K. Agoston,* Algebraic Topology (1976)
33. *K. R. Goodearl,* Ring Theory (1976)
34. *L. E. Mansfield,* Linear Algebra with Geometric Applications (1976)
35. *N. J. Pullman,* Matrix Theory and Its Applications (1976)
36. *B. R. McDonald,* Geometric Algebra Over Local Rings (1976)
37. *C. W. Groetsch,* Generalized Inverses of Linear Operators (1977)
38. *J. E. Kuczkowski and J. L. Gersting,* Abstract Algebra (1977)
39. *C. O. Christenson and W. L. Voxman,* Aspects of Topology (1977)
40. *M. Nagata,* Field Theory (1977)
41. *R. L. Long,* Algebraic Number Theory (1977)
42. *W. F. Pfeffer,* Integrals and Measures (1977)
43. *R. L. Wheeden and A. Zygmund,* Measure and Integral (1977)
44. *J. H. Curtiss,* Introduction to Functions of a Complex Variable (1978)
45. *K. Hrbacek and T. Jech,* Introduction to Set Theory (1978)
46. *W. S. Massey,* Homology and Cohomology Theory (1978)
47. *M. Marcus,* Introduction to Modern Algebra (1978)
48. *E. C. Young,* Vector and Tensor Analysis (1978)
49. *S. B. Nadler, Jr.,* Hyperspaces of Sets (1978)
50. *S. K. Segal,* Topics in Group Kings (1978)
51. *A. C. M. van Rooij,* Non-Archimedean Functional Analysis (1978)
52. *L. Corwin and R. Szczarba,* Calculus in Vector Spaces (1979)
53. *C. Sadosky,* Interpolation of Operators and Singular Integrals (1979)
54. *J. Cronin,* Differential Equations (1980)
55. *C. W. Groetsch,* Elements of Applicable Functional Analysis (1980)

56. *I. Vaisman*, Foundations of Three-Dimensional Euclidean Geometry (1980)
57. *H. I. Freedan*, Deterministic Mathematical Models in Population Ecology (1980)
58. *S. B. Chae*, Lebesgue Integration (1980)
59. *C. S. Rees et al.*, Theory and Applications of Fourier Analysis (1981)
60. *L. Nachbin*, Introduction to Functional Analysis (R. M. Aron, trans.) (1981)
61. *G. Orzech and M. Orzech*, Plane Algebraic Curves (1981)
62. *R. Johnsonbaugh and W. E. Pfaffenberger*, Foundations of Mathematical Analysis (1981)
63. *W. L. Voxman and R. H. Goetschel*, Advanced Calculus (1981)
64. *L. J. Corwin and R. H. Szczarba*, Multivariable Calculus (1982)
65. *V. I. Istrǎţescu*, Introduction to Linear Operator Theory (1981)
66. *R. D. Järvinen*, Finite and Infinite Dimensional Linear Spaces (1981)
67. *J. K. Beem and P. E. Ehrlich*, Global Lorentzian Geometry (1981)
68. *D. L. Armacost*, The Structure of Locally Compact Abelian Groups (1981)
69. *J. W. Brewer and M. K. Smith, eds.*, Emmy Noether: A Tribute (1981)
70. *K. H. Kim*, Boolean Matrix Theory and Applications (1982)
71. *T. W. Wieting*, The Mathematical Theory of Chromatic Plane Ornaments (1982)
72. *D. B. Gauld*, Differential Topology (1982)
73. *R. L. Faber*, Foundations of Euclidean and Non-Euclidean Geometry (1983)
74. *M. Carmeli*, Statistical Theory and Random Matrices (1983)
75. *J. H. Carruth et al.*, The Theory of Topological Semigroups (1983)
76. *R. L. Faber*, Differential Geometry and Relativity Theory (1983)
77. *S. Barnett*, Polynomials and Linear Control Systems (1983)
78. *G. Karpilovsky*, Commutative Group Algebras (1983)
79. *F. Van Oystaeyen and A. Verschoren*, Relative Invariants of Rings (1983)
80. *I. Vaisman*, A First Course in Differential Geometry (1984)
81. *G. W. Swan*, Applications of Optimal Control Theory in Biomedicine (1984)
82. *T. Petrie and J. D. Randall*, Transformation Groups on Manifolds (1984)
83. *K. Goebel and S. Reich*, Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings (1984)
84. *T. Albu and C. Năstăsescu*, Relative Finiteness in Module Theory (1984)
85. *K. Hrbacek and T. Jech*, Introduction to Set Theory: Second Edition (1984)
86. *F. Van Oystaeyen and A. Verschoren*, Relative Invariants of Rings (1984)
87. *B. R. McDonald*, Linear Algebra Over Commutative Rings (1984)
88. *M. Namba*, Geometry of Projective Algebraic Curves (1984)
89. *G. F. Webb*, Theory of Nonlinear Age-Dependent Population Dynamics (1985)
90. *M. R. Bremner et al.*, Tables of Dominant Weight Multiplicities for Representations of Simple Lie Algebras (1985)
91. *A. E. Fekete*, Real Linear Algebra (1985)
92. *S. B. Chae*, Holomorphy and Calculus in Normed Spaces (1985)
93. *A. J. Jerri*, Introduction to Integral Equations with Applications (1985)
94. *G. Karpilovsky*, Projective Representations of Finite Groups (1985)
95. *L. Narici and E. Beckenstein*, Topological Vector Spaces (1985)
96. *J. Weeks*, The Shape of Space (1985)
97. *P. R. Gribik and K. O. Kortanek*, Extremal Methods of Operations Research (1985)
98. *J.-A. Chao and W. A. Woyczynski, eds.*, Probability Theory and Harmonic Analysis (1986)
99. *G. D. Crown et al.*, Abstract Algebra (1986)
100. *J. H. Carruth et al.*, The Theory of Topological Semigroups, Volume 2 (1986)
101. *R. S. Doran and V. A. Belfi*, Characterizations of C^*-Algebras (1986)
102. *M. W. Jeter*, Mathematical Programming (1986)
103. *M. Altman*, A Unified Theory of Nonlinear Operator and Evolution Equations with Applications (1986)
104. *A. Verschoren*, Relative Invariants of Sheaves (1987)
105. *R. A. Usmani*, Applied Linear Algebra (1987)
106. *P. Blass and J. Lang*, Zariski Surfaces and Differential Equations in Characteristic $p > 0$ (1987)
107. *J. A. Reneke et al.*, Structured Hereditary Systems (1987)
108. *H. Busemann and B. B. Phadke*, Spaces with Distinguished Geodesics (1987)
109. *R. Harte*, Invertibility and Singularity for Bounded Linear Operators (1988)
110. *G. S. Ladde et al.*, Oscillation Theory of Differential Equations with Deviating Arguments (1987)
111. *L. Dudkin et al.*, Iterative Aggregation Theory (1987)
112. *T. Okubo*, Differential Geometry (1987)

113. *D. L. Stancl and M. L. Stancl,* Real Analysis with Point-Set Topology (1987)
114. *T. C. Gard,* Introduction to Stochastic Differential Equations (1988)
115. *S. S. Abhyankar,* Enumerative Combinatorics of Young Tableaux (1988)
116. *H. Strade and R. Farnsteiner,* Modular Lie Algebras and Their Representations (1988)
117. *J. A. Huckaba,* Commutative Rings with Zero Divisors (1988)
118. *W. D. Wallis,* Combinatorial Designs (1988)
119. *W. Więsław,* Topological Fields (1988)
120. *G. Karpilovsky,* Field Theory (1988)
121. *S. Caenepeel and F. Van Oystaeyen,* Brauer Groups and the Cohomology of Graded Rings (1989)
122. *W. Kozlowski,* Modular Function Spaces (1988)
123. *E. Lowen-Colebunders,* Function Classes of Cauchy Continuous Maps (1989)
124. *M. Pavel,* Fundamentals of Pattern Recognition (1989)
125. *V. Lakshmikantham et al.,* Stability Analysis of Nonlinear Systems (1989)
126. *R. Sivaramakrishnan,* The Classical Theory of Arithmetic Functions (1989)
127. *N. A. Watson,* Parabolic Equations on an Infinite Strip (1989)
128. *K. J. Hastings,* Introduction to the Mathematics of Operations Research (1989)
129. *B. Fine,* Algebraic Theory of the Bianchi Groups (1989)
130. *D. N. Dikranjan et al.,* Topological Groups (1989)
131. *J. C. Morgan II,* Point Set Theory (1990)
132. *P. Biler and A. Witkowski,* Problems in Mathematical Analysis (1990)
133. *H. J. Sussmann,* Nonlinear Controllability and Optimal Control (1990)
134. *J.-P. Florens et al.,* Elements of Bayesian Statistics (1990)
135. *N. Shell,* Topological Fields and Near Valuations (1990)
136. *B. F. Doolin and C. F. Martin,* Introduction to Differential Geometry for Engineers (1990)
137. *S. S. Holland, Jr.,* Applied Analysis by the Hilbert Space Method (1990)
138. *J. Okniński,* Semigroup Algebras (1990)
139. *K. Zhu,* Operator Theory in Function Spaces (1990)
140. *G. B. Price,* An Introduction to Multicomplex Spaces and Functions (1991)
141. *R. B. Darst,* Introduction to Linear Programming (1991)
142. *P. L. Sachdev,* Nonlinear Ordinary Differential Equations and Their Applications (1991)
143. *T. Husain,* Orthogonal Schauder Bases (1991)
144. *J. Foran,* Fundamentals of Real Analysis (1991)
145. *W. C. Brown,* Matrices and Vector Spaces (1991)
146. *M. M. Rao and Z. D. Ren,* Theory of Orlicz Spaces (1991)
147. *J. S. Golan and T. Head,* Modules and the Structures of Rings (1991)
148. *C. Small,* Arithmetic of Finite Fields (1991)
149. *K. Yang,* Complex Algebraic Geometry (1991)
150. *D. G. Hoffman et al.,* Coding Theory (1991)
151. *M. O. González,* Classical Complex Analysis (1992)
152. *M. O. González,* Complex Analysis (1992)
153. *L. W. Baggett,* Functional Analysis (1992)
154. *M. Sniedovich,* Dynamic Programming (1992)
155. *R. P. Agarwal,* Difference Equations and Inequalities (1992)
156. *C. Brezinski,* Biorthogonality and Its Applications to Numerical Analysis (1992)
157. *C. Swartz,* An Introduction to Functional Analysis (1992)
158. *S. B. Nadler, Jr.,* Continuum Theory (1992)
159. *M. A. Al-Gwaiz,* Theory of Distributions (1992)
160. *E. Perry,* Geometry: Axiomatic Developments with Problem Solving (1992)
161. *E. Castillo and M. R. Ruiz-Cobo,* Functional Equations and Modelling in Science and Engineering (1992)
162. *A. J. Jerri,* Integral and Discrete Transforms with Applications and Error Analysis (1992)
163. *A. Charlier et al.,* Tensors and the Clifford Algebra (1992)
164. *P. Biler and T. Nadzieja,* Problems and Examples in Differential Equations (1992)
165. *E. Hansen,* Global Optimization Using Interval Analysis (1992)
166. *S. Guerre-Delabrière,* Classical Sequences in Banach Spaces (1992)
167. *Y. C. Wong,* Introductory Theory of Topological Vector Spaces (1992)
168. *S. H. Kulkarni and B. V. Limaye,* Real Function Algebras (1992)
169. *W. C. Brown,* Matrices Over Commutative Rings (1993)
170. *J. Loustau and M. Dillon,* Linear Geometry with Computer Graphics (1993)
171. *W. V. Petryshyn,* Approximation-Solvability of Nonlinear Functional and Differential Equations (1993)

172. *E. C. Young*, Vector and Tensor Analysis: Second Edition (1993)
173. *T. A. Bick*, Elementary Boundary Value Problems (1993)
174. *M. Pavel*, Fundamentals of Pattern Recognition: Second Edition (1993)
175. *S. A. Albeverio et al.*, Noncommutative Distributions (1993)
176. *W. Fulks*, Complex Variables (1993)
177. *M. M. Rao*, Conditional Measures and Applications (1993)
178. *A. Janicki and A. Weron*, Simulation and Chaotic Behavior of α-Stable Stochastic Processes (1994)
179. *P. Neittaanmäki and D. Tiba*, Optimal Control of Nonlinear Parabolic Systems (1994)
180. *J. Cronin*, Differential Equations: Introduction and Qualitative Theory, Second Edition (1994)
181. *S. Heikkilä and V. Lakshmikantham*, Monotone Iterative Techniques for Discontinuous Nonlinear Differential Equations (1994)
182. *X. Mao*, Exponential Stability of Stochastic Differential Equations (1994)
183. *B. S. Thomson*, Symmetric Properties of Real Functions (1994)
184. *J. E. Rubio*, Optimization and Nonstandard Analysis (1994)
185. *J. L. Bueso et al.*, Compatibility, Stability, and Sheaves (1995)
186. *A. N. Michel and K. Wang*, Qualitative Theory of Dynamical Systems (1995)
187. *M. R. Darnel*, Theory of Lattice-Ordered Groups (1995)
188. *Z. Naniewicz and P. D. Panagiotopoulos*, Mathematical Theory of Hemivariational Inequalities and Applications (1995)
189. *L. J. Corwin and R. H. Szczarba*, Calculus in Vector Spaces: Second Edition (1995)
190. *L. H. Erbe et al.*, Oscillation Theory for Functional Differential Equations (1995)
191. *S. Agaian et al.*, Binary Polynomial Transforms and Nonlinear Digital Filters (1995)
192. *M. I. Gil'*, Norm Estimations for Operation-Valued Functions and Applications (1995)
193. *P. A. Grillet*, Semigroups: An Introduction to the Structure Theory (1995)
194. *S. Kichenassamy*, Nonlinear Wave Equations (1996)
195. *V. F. Krotov*, Global Methods in Optimal Control Theory (1996)
196. *K. I. Beidar et al.*, Rings with Generalized Identities (1996)
197. *V. I. Arnautov et al.*, Introduction to the Theory of Topological Rings and Modules (1996)
198. *G. Sierksma*, Linear and Integer Programming (1996)
199. *R. Lasser*, Introduction to Fourier Series (1996)
200. *V. Sima*, Algorithms for Linear-Quadratic Optimization (1996)
201. *D. Redmond*, Number Theory (1996)
202. *J. K. Beem et al.*, Global Lorentzian Geometry: Second Edition (1996)
203. *M. Fontana et al.*, Prüfer Domains (1997)
204. *H. Tanabe*, Functional Analytic Methods for Partial Differential Equations (1997)
205. *C. Q. Zhang*, Integer Flows and Cycle Covers of Graphs (1997)
206. *E. Spiegel and C. J. O'Donnell*, Incidence Algebras (1997)
207. *B. Jakubczyk and W. Respondek*, Geometry of Feedback and Optimal Control (1998)
208. *T. W. Haynes et al.*, Fundamentals of Domination in Graphs (1998)
209. *T. W. Haynes et al.*, Domination in Graphs: Advanced Topics (1998)
210. *L. A. D'Alotto et al.*, A Unified Signal Algebra Approach to Two-Dimensional Parallel Digital Signal Processing (1998)
211. *F. Halter-Koch*, Ideal Systems (1998)
212. *N. K. Govil et al.*, Approximation Theory (1998)
213. *R. Cross*, Multivalued Linear Operators (1998)
214. *A. A. Martynyuk*, Stability by Liapunov's Matrix Function Method with Applications (1998)
215. *A. Favini and A. Yagi*, Degenerate Differential Equations in Banach Spaces (1999)
216. *A. Illanes and S. Nadler, Jr.*, Hyperspaces: Fundamentals and Recent Advances (1999)
217. *G. Kato and D. Struppa*, Fundamentals of Algebraic Microlocal Analysis (1999)
218. *G. X.-Z. Yuan*, KKM Theory and Applications in Nonlinear Analysis (1999)
219. *D. Motreanu and N. H. Pavel*, Tangency, Flow Invariance for Differential Equations, and Optimization Problems (1999)
220. *K. Hrbacek and T. Jech*, Introduction to Set Theory, Third Edition (1999)
221. *G. E. Kolosov*, Optimal Design of Control Systems (1999)
222. *N. L. Johnson*, Subplane Covered Nets (2000)
223. *B. Fine and G. Rosenberger*, Algebraic Generalizations of Discrete Groups (1999)
224. *M. Väth*, Volterra and Integral Equations of Vector Functions (2000)
225. *S. S. Miller and P. T. Mocanu*, Differential Subordinations (2000)

226. *R. Li et al.*, Generalized Difference Methods for Differential Equations: Numerical Analysis of Finite Volume Methods (2000)
227. *H. Li and F. Van Oystaeyen*, A Primer of Algebraic Geometry (2000)
228. *R. P. Agarwal*, Difference Equations and Inequalities: Theory, Methods, and Applications, Second Edition (2000)
229. *A. B. Kharazishvili*, Strange Functions in Real Analysis (2000)
230. *J. M. Appell et al.*, Partial Integral Operators and Integro-Differential Equations (2000)
231. *A. I. Prilepko et al.*, Methods for Solving Inverse Problems in Mathematical Physics (2000)
232. *F. Van Oystaeyen*, Algebraic Geometry for Associative Algebras (2000)
233. *D. L. Jagerman*, Difference Equations with Applications to Queues (2000)
234. *D. R. Hankerson et al.*, Coding Theory and Cryptography: The Essentials, Second Edition, Revised and Expanded (2000)
235. *S. Dăscălescu et al.*, Hopf Algebras: An Introduction (2001)
236. *R. Hagen et al.*, C*-Algebras and Numerical Analysis (2001)
237. *Y. Talpaert*, Differential Geometry: With Applications to Mechanics and Physics (2001)
238. *R. H. Villarreal*, Monomial Algebras (2001)
239. *A. N. Michel et al.*, Qualitative Theory of Dynamical Systems, Second Edition (2001)
240. *A. A. Samarskii*, The Theory of Difference Schemes (2001)
241. *J. Knopfmacher and W.-B. Zhang*, Number Theory Arising from Finite Fields (2001)
242. *S. Leader*, The Kurzweil-Henstock Integral and Its Differentials (2001)
243. *M. Biliotti et al.*, Foundations of Translation Planes (2001)
244. *A. N. Kochubei*, Pseudo-Differential Equations and Stochastics over Non-Archimedean Fields (2001)
245. *G. Sierksma*, Linear and Integer Programming, Second Edition (2002)
246. *A. A. Martynyuk*, Qualitative Methods in Nonlinear Dynamics: Novel Approaches to Liapunov's Matrix Functions (2002)
247. *B. G. Pachpatte*, Inequalities for Finite Difference Equations (2002)
248. *A. N. Michel and D. Liu*, Qualitative Analysis and Synthesis of Recurrent Neural Networks (2002)
249. *J. Weeks*, The Shape of Space (2002)

Additional Volumes in Preparation

This Page Intentionally Left Blank

Part I

Surfaces and Three-Manifolds

This Page Intentionally Left Blank

1

Flatland

In 1884 an extraordinary individual named A Square succeeded in publishing his memoirs. Actually an intermediary by the name of Abbott published them for him—A Square himself was in prison for heresy at the time. A Square was extraordinary not because he had such an odd name, but rather because he had such a descriptive and accurate name. For you see, A Square was a square.

Now you might be wondering just where A Square lived. After all, you wouldn't expect to find a two-dimensional square living in a three-dimensional uni-

verse such as ours. You might allow for a slightly thickened square, say a creature with the dimensions of a sheet of paper, but certainly not a completely flat individual like A Square. Anyhow, A Square didn't live in our three-dimensional universe. He lived in Flatland, a two-dimensional universe resembling a giant plane.

Flatland also happens to be the title under which A Square's memoirs were published. It's now available in paperback, and I recommend it highly. In 1907, C. H. Hinton published a similar book, *An Episode of Flatland*. The chief difference between these books is that the residents of Flatland proper can move freely about their two-dimensional universe, whereas the inhabitants of Hinton's world are constrained by gravity to living on the circular edge of their disk-shaped planet Astria (Figure 1.1). For the full story on the lore of Astria, see A. K. Dewdney's *The Planiverse*.

Getting back to the subject at hand, the Flatlanders all thought that Flatland was a giant plane, what we Spacelanders would call a Euclidean plane. To be accurate, I should say that they *assumed* that Flatland was a plane, since nobody ever gave the issue any thought. Well, almost nobody. Once a physicist by the name of A Stone had proposed an alternative theory, something about Flatland having a finite area, yet having no boundary. He compared Flatland to a circle. For the most part people didn't understand him. It was obvious that a circle had a finite circumference and no endpoints, but what did that have to do with

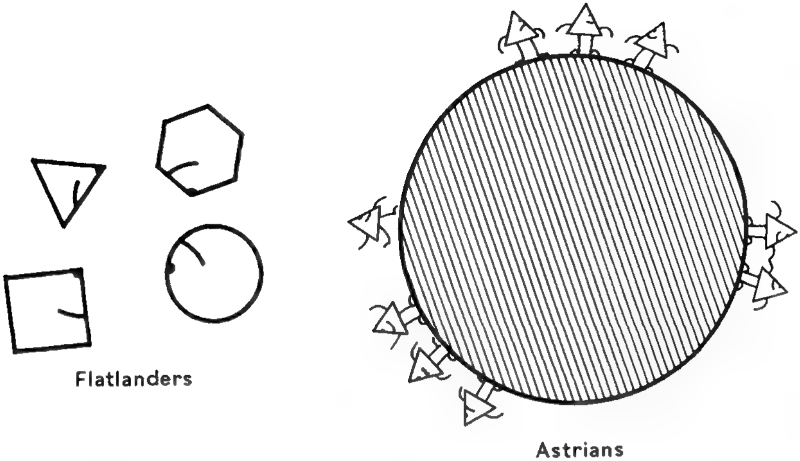


Figure 1.1 Flatlanders move freely in a "plane," while Astrians are confined to the edge of their disk-shaped planet Astria.

Flatland, which obviously had an infinite area? At least part of the problem was linguistic: The only word for "plane" was the word for "Flatland" itself, so to express the idea that Flatland was *not* a plane, one was trapped into stating that "Flatland is not Flatland." Needless to say, this theory attracted few disciples.

A Square, though, was among the few. He was particularly intrigued by the idea that a person could set out in one direction and come back from the opposite direction, without ever having turned around. He was so intrigued that he wanted to try it out. The Flatlanders were for the most part a timid lot, and few had ever traveled more than a day or two's jour-

ney beyond the outlying farms of Flatsburgh. A Square reasoned that if he were willing to spend a month tromping eastward through the woods, he might just have a shot at coming back from the west.

He was delighted when two friends volunteered to go with him. The friends, A Pentagon and A Hexagon, didn't believe any of A Square's theories—they just wanted to keep him out of trouble. To this end they insisted that A Square buy up all the red thread he could find in Flatsburgh. The idea was that they would lay out a trail of red thread behind them, so that after they had traveled for a month and given up, they could then find their way back to Flatsburgh.

As it turned out, the thread was unnecessary. Much to A Square's delight—and A Pentagon's and A Hexagon's relief—they returned from the west after three weeks of travel. Not that this convinced anyone of anything. Even A Pentagon and A Hexagon thought that they must have veered slightly to one side or the other, bending their route into a giant circle in the plane of Flatland (Figure 1.2). A Square had no reply to their theory, but this did little to dampen his enthusiasm. He was ready to try it again!

By now red thread was in short supply in Flatland, so this time A Square laid out a trail of blue thread to mark his route. He set out to the north, and, sure enough, returned two weeks later from the south. Again everyone assumed that he had simply veered in a circle, and counted him lucky for getting back at all.

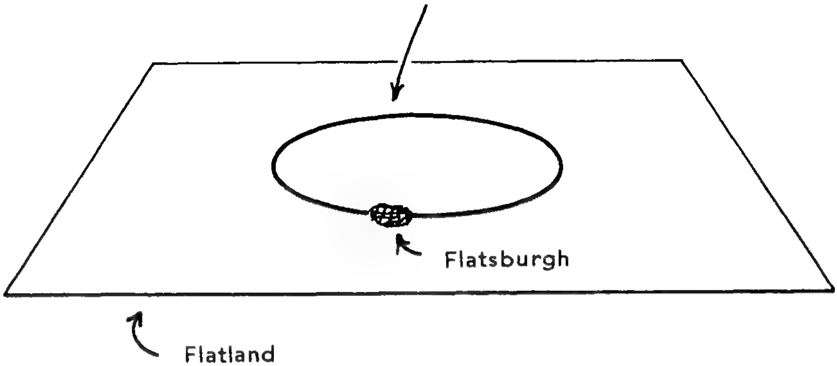


Figure 1.2 Even A Square's companions thought they had veered in a circle.

A Square was mystified that his journey was so much shorter this time, but something else bothered him even more: he had never come across the red thread they laid out on the first journey. The physicists of Flatland were equally intrigued. They confirmed that even if Flatland were a so-called "hypercircle" as A Stone had suggested, the two threads would still cross (Figure 1.3). There was, of course, the possibility that the red thread had broken for one reason or another. To investigate this possibility, the scientists formed two expeditions: one party retraced the red thread, the other retraced the blue. Both threads were found to be intact.

The Mystery of the Nonintersecting Threads remained a mystery for quite a few years. Some of the bolder Flatlanders even took to retracing the threads periodically as a sort of pilgrimage. The first hint of a

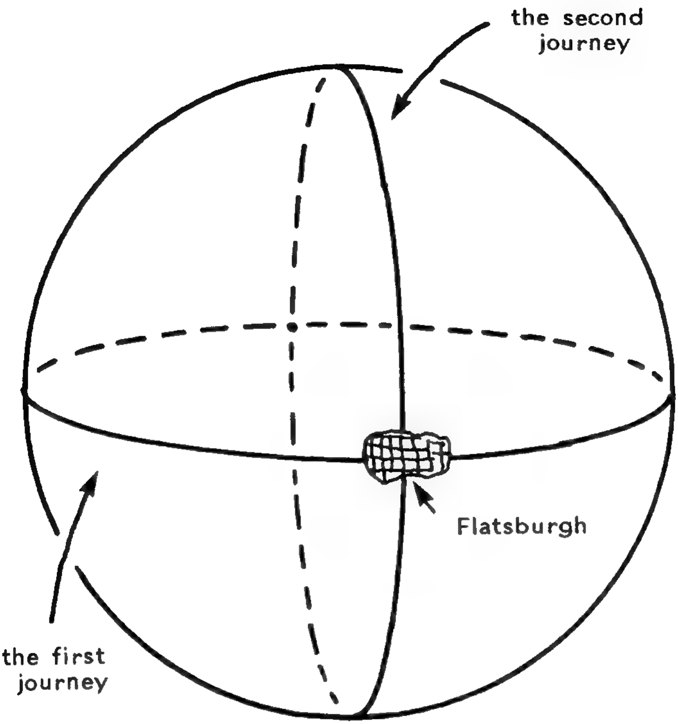


Figure 1.3 The two threads ought to cross, even if Flatland were a "hypercircle" (i.e., a sphere).

resolution came when a physicist proposed that Flatland should be regarded neither as a "Flatland" (i.e., a plane) nor as a hypercircle, but as something he called a "torus." At first no one had any idea what he was talking about. Gradually though, people agreed that this theory resolved the Mystery of the Nonintersecting Threads, and everyone was happy about that. So for many years Flatland was thought to be a torus (Figure 1.4).

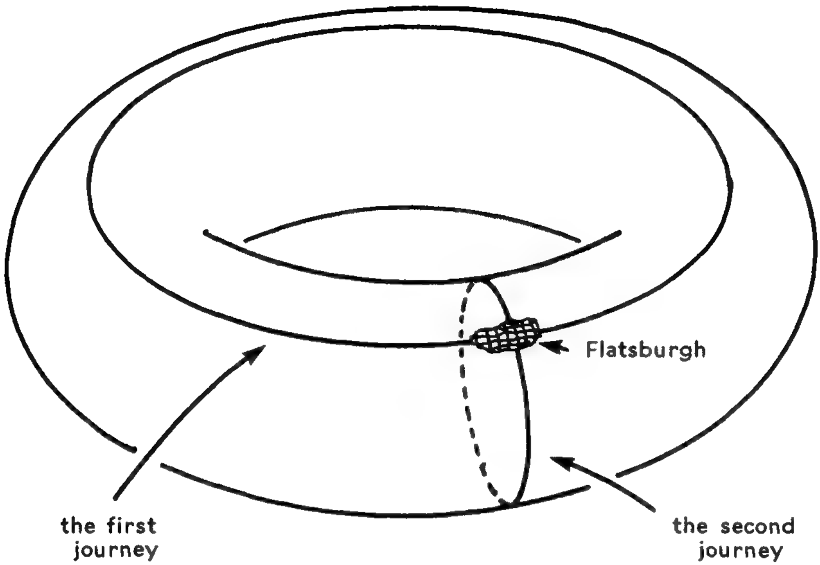


Figure 1.4 Spacelanders sometimes visualize a torus as the surface of a doughnut.

Until one day somebody came up with yet another theory on the "shape" of Flatland. This theory explained the Mystery of the Nonintersecting Threads just as well as the torus theory did, but it gave a different overall view of Flatland.

And this new theory was just the first of many. For the next few months people were constantly coming up with new possibilities for the shape of Flatland (Figure 1.5). Soon a vast Universal Survey was undertaken to map all of Flatland and thereby determine its true shape once and for all . . .

(*STORY TO BE CONTINUED IN CHAPTER 4*)

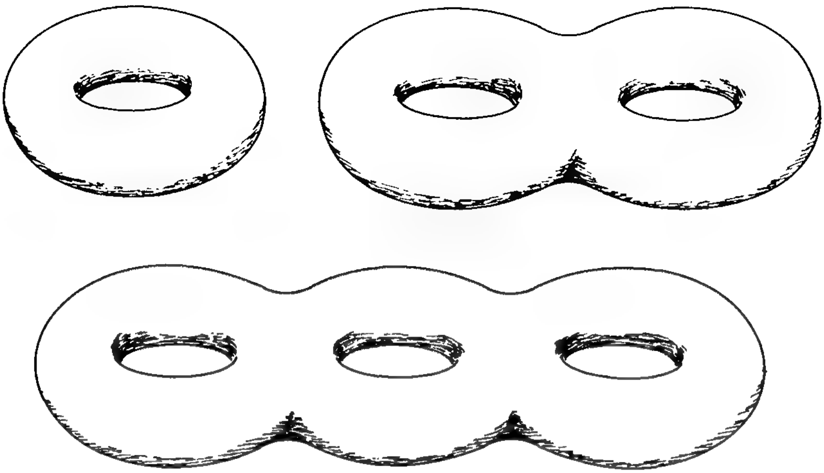


Figure 1.5 Some possible shapes for Flatland.

As Spacelanders we have three dimensions available for drawing pictures like Figure 1.5, so it's easy for us to understand how a two-dimensional universe can close back on itself. The Flatlanders inhabiting such a universe would have a much tougher time. To sympathize with their feelings, try imagining yourself in each of the following situations.[1]

1. You are on an expedition to a distant galaxy in search of intelligent life. When you reach the galaxy you head for the most hospitable-looking planet you can find, only to discovery that you're back on Earth.

---

[1]Warning: These situations are designed to stimulate the imagination. Don't worry about technical complications!

2. You are an astronomer. You seem to be observing the exact same object in two different locations in the sky.
3. You are a radio astronomer searching for signals from extraterrestrials. You have detected a faint signal coming from a distant galaxy. Once you tune it in you recognize it as a broadcast of the old TV show "Father Knows Best."

Each of the above situations leads you to suspect that space is built differently than you thought it was. That is, space seems to have a different shape than the obvious one you assumed it had. Not that you have any idea what this actual shape is!

In fact, no one knows what the shape of the real universe is. But people do know a fair amount about what the *possible* shapes are. These possible shapes are the topic of this book. Such a possible shape is called a three-dimensional manifold, or three-manifold for short. (Similarly, a two-dimensional shape for Flatland is called a two-dimensional manifold, or, more commonly, a surface.) At this point your conception of a three-manifold is probably pretty vague. Don't worry: we'll start seeing some examples in Chapter 2. The main thing now is to realize that our universe might conceivably close back on itself, just as the various surfaces representing Flatland close back on themselves.

This book centers on a series of examples of three-manifolds. Rather than developing an extensive theory of these manifolds, you'll come to know each of them in a visual and intuitive way. Obviously this is not an easy task. Imagine the difficulties A Square would have in communicating to A Hexagon the true nature of a torus. A Square cannot draw a definitive picture of a torus, being confined to two dimensions as he is. Similarly, we cannot draw a definitive picture of any three-manifold.

There is some hope, though. You can use tricks to define various three-manifolds, and as you work with them over a period of time you'll find your intuition for them growing steadily. The human mind is remarkably flexible in this regard. *Just be sure to read slowly and give things plenty of time to digest*. At most a chapter, and often as little as a single exercise, will be plenty for one sitting.

This book provides not a series of answers, but rather a series of questions designed to lead you to your own intuitive understanding of three-manifolds. Prepare your imagination for a workout!

2

Gluing

A popular video game pits two players in biplanes in aerial combat on a TV screen. An interesting feature of the game is that when a biplane flies off one edge of the screen it doesn't crash, but rather it comes back from the opposite edge of the screen (Figure 2.1). Mathematically speaking, the screen's edges have been "glued" together. (The gluing is purely abstract: there is no need to *physically* connect the edges.) A square or rectangle whose opposite edges are abstractly glued in this fashion is called a torus or, more precisely, a flat two-dimensional torus. There is a con-

nection between this flat two-dimensional torus and the doughnut-surface torus of Chapter 1, but for the time being *you should forget the doughnut surface entirely*.

You can play interactive torus games online at *www.northnet.org/weeks/SoS*.

Exercise 2.1 Play a few games of torus tic-tac-toe with a willing friend. The rules are the same as in traditional tic-tac-toe, except here the opposite sides of the board are glued to form a torus, just as the

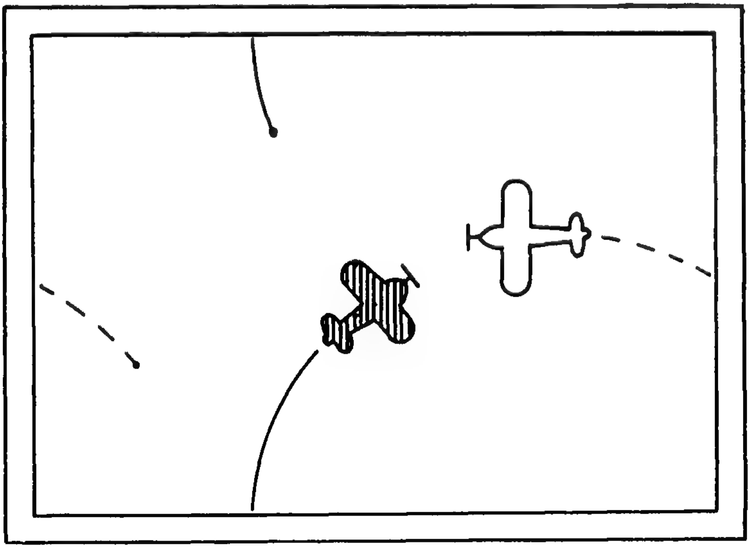


Figure 2.1 The biplanes can fly across the edges of the TV screen.

opposite sides of the TV screen are glued in the biplane game. So, for example, the three Xs in Figure 2.2 constitute a winning three-in-a-row. □

Exercise 2.2 Figure 2.3 shows a torus tic-tac-toe game in progress. It's X's turn. Where should he move? If it were O's turn instead, where would his best move be? (*Answers to exercises are found in Appendix A in the back of the book.*) □

The positions shown in Figure 2.4 are all equivalent in torus tic-tac-toe. The second position is obtained from the first by moving everything "up" one notch (when the top row moves "up" it naturally reap-

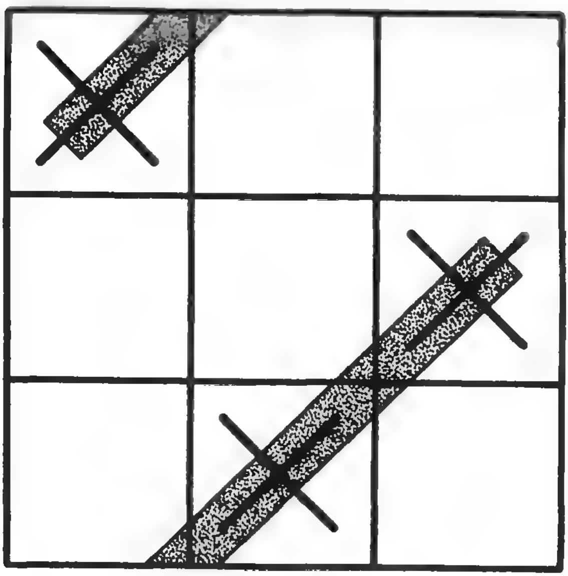


Figure 2.2 These Xs are three-in-a-row if the board is imagined to represent a torus.

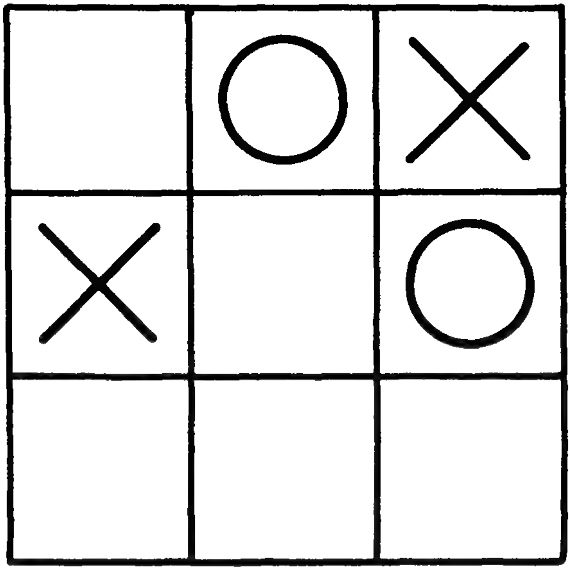


Figure 2.3 What is X's best move? What is O's best move?

pears at the bottom). Similarly, the third position is the result of moving everything in the second position one notch to the right. The fourth position is obtained a little differently: it results from rotating the third position one quarter turn clockwise.

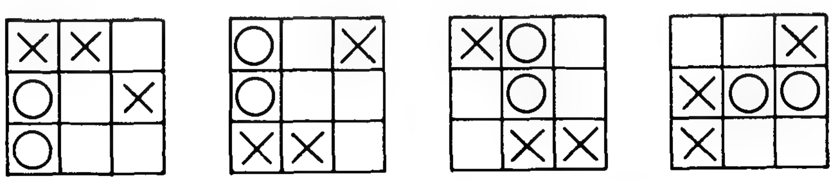


Figure 2.4 These four positions are equivalent in torus tic-tac-toe.

Exercise 2.3 Which of the positions in Figure 2.5 are equivalent in torus tic-tac-toe? □

Exercise 2.4 In torus tic-tac-toe, how many essentially different opening moves does the first player have? How many different responses does her opponent have? Is either player guaranteed to win, assuming optimal play? (In traditional tic-tac-toe, either player can guarantee a draw.) □

Exercise 2.5 Chess on a torus is more challenging than tic-tac-toe. Consider the position shown in Figure 2.6. Which black pieces does the white knight threaten? Which black pieces threaten it? □

Exercise 2.6 Figure 2.7 shows some pieces on a torus chessboard. Which black pieces are threatened by *both* the white knight and the white queen? □

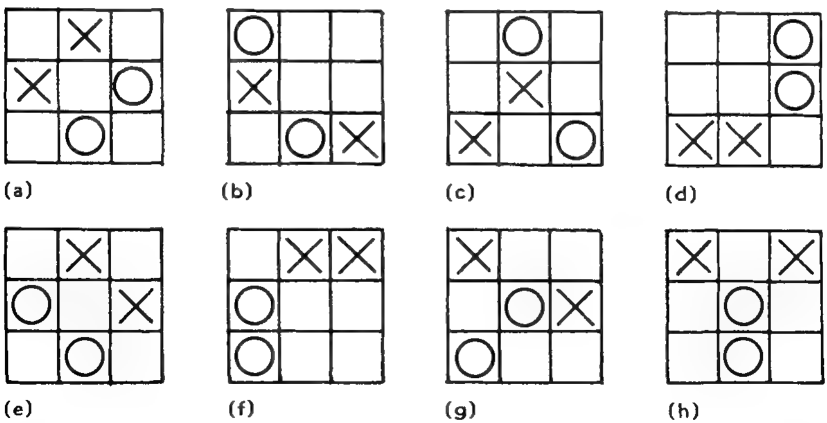


Figure 2.5 Which of these positions are equivalent in torus tic-tac-toe?

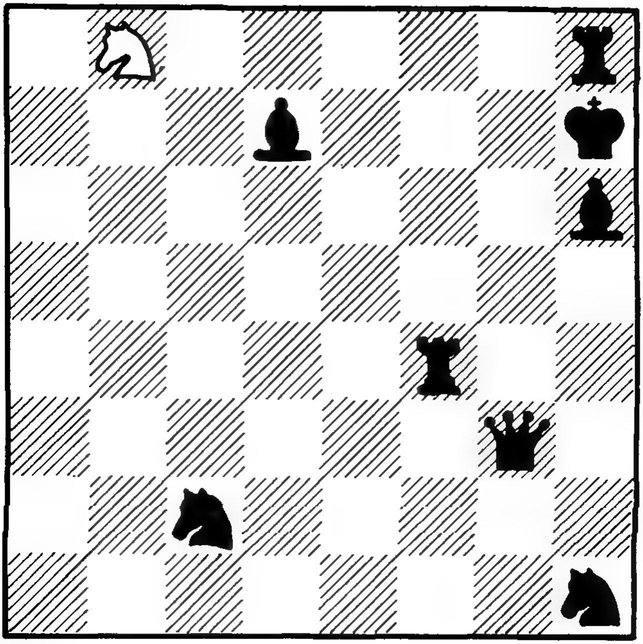


Figure 2.6 A chessboard becomes a torus when opposite sides are connected.

Exercise 2.7 Find a friend and play a few games of torus chess. The usual starting position just won't do for torus chess (try it and you'll see why). Instead either use the starting position of Figure 2.8, or make up a starting position of your own. All the pieces move normally except the pawns: a pawn moves one space forward, backward, to the left or to the right, and captures by moving one space on any diagonal. □

Exercise 2.8 When a bishop goes out the upper right-hand corner of a torus chessboard, where does he return? □

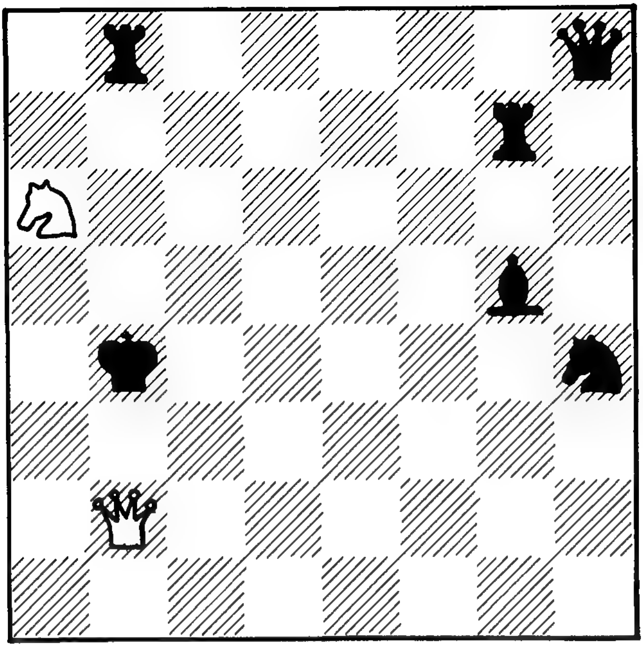


Figure 2.7 Which black pieces are threatened by *both* the white knight and the white queen?

Exercise 2.9 In torus chess, can a knight and a bishop simultaneously threaten each other? □

The flat torus we've been using for tic-tac-toe and chess is a two-dimensional manifold, just like the two-dimensional manifolds in Figure 1.5. Like the two-manifolds of Figure 1.5, the flat torus has a finite area and no edges. *Unlike* those two-manifolds, the flat torus is defined abstractly—via gluing—instead of being drawn in three-dimensional space. The same trick works to define three-dimensional manifolds without resorting to pictures in four-dimensional space.

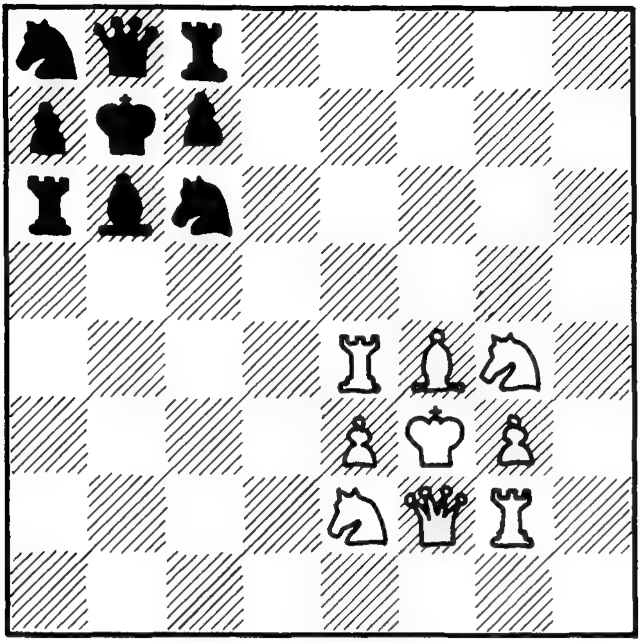


Figure 2.8 Here's one possible starting position for torus chess.

Our first three-dimensional manifold is analogous to the flat two-dimensional torus. It's called a three-dimensional torus, or three-torus for short. To construct it, start with a solid block of space—the room you're in will do fine just so it's rectangular. Imagine the left wall glued to the right wall, not in the sense that you'd physically carry out the gluing, but in the sense that if you walked through the left wall you'd find yourself emerging from the right wall. Imagine the front wall glued to the back wall, and the floor

glued to the ceiling, in the same manner. You are now sitting in a three-dimensional torus! This three-torus has no edges, and its total volume is just the volume of the room you started with.

Exercise 2.10 What do you see when you look through the "wall" of the three-torus described above? For that matter, what do you see when you look through the "floor" or the "ceiling"? □

Imagine a three-torus made from a cube ten meters on a side. The cube contains a jungle gym consisting of a rectangular lattice of pipes (Figure 2.9); each segment of pipe is one meter long. When the cube's faces are glued to form a three-torus, the jungle gym continues uniformly across each face.

It would be fairly boring playing in this jungle gym by yourself. Sure, you could climb up a few meters or over a few meters, but your new location would be just like your old one. The visual effects would be interesting, though. You could look up ten meters, or over ten meters, and see yourself. You couldn't go meet yourself, though, since as soon as you started climbing towards your "other self," she would start climbing away from you in a (futile) effort to go meet *her* other self ten meters further away!

All in all, playing in this jungle gym would be a lot more fun with a friend. For example, your friend could wait for you while you climbed up ten meters to meet him. Of course, while you were climbing *up*, he

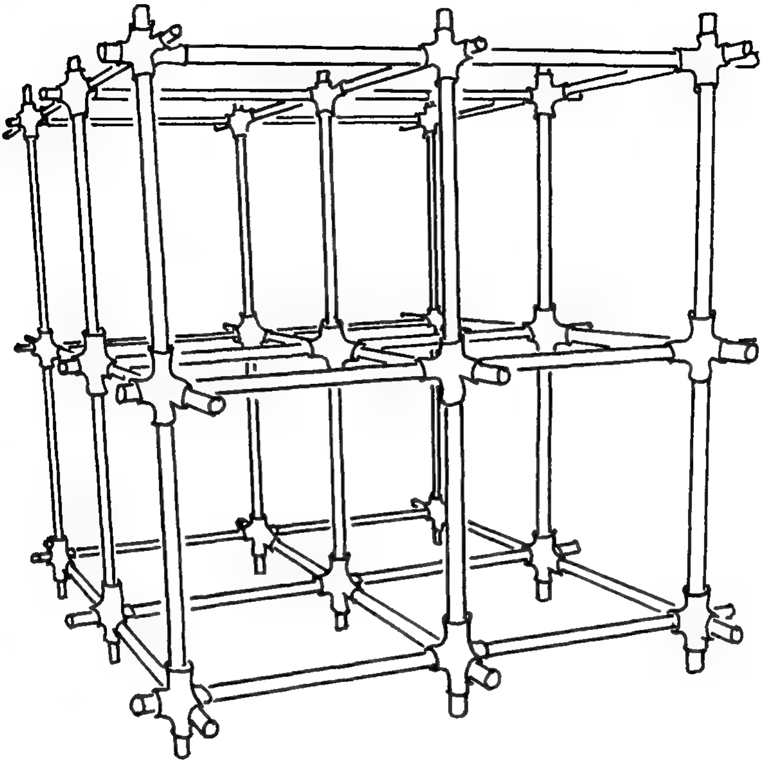


Figure 2.9 A section of jungle gym.

could crawl ten meters *over*, so that when you got to the rendezvous point he would be coming in from the side rather than just waiting there. Tag wold be especially fun in a three-torus.

Exercise 2.11 One's imagination can go wild in a three-torus. For example, you could imagine flying around in real, three-dimensional biplanes. Or you could imagine a completely urbanized three-torus in which all north–south streets are one way north-

bound, and all east–west streets are one way eastbound, and all elevators go only up (you could still get wherever you wanted to go). Imagine other things you could do in a three-torus. □

Exercise 2.12 How could you play catch by yourself in a three-torus? □

In theory, if the universe is a three-torus we should be able to look out into space and see ourselves. Does the fact that astronomers have not done so mean that the real universe cannot be a three-dimensional torus? Not at all! The universe is only 10 or 20 billion years old, so if it were a very large three-dimensional torus—say 60 billion light-years across at its present stage of evolution—then no light would yet have had enough time to make a complete trip across. Another possibility is that we *are* in fact seeing all the way across the universe, but we just don't know it: when we look off into distant space we see things as they were billions of years ago, and billions of years ago our galaxy looked different than it does now. (This effect occurs because the light that enters a telescope today left its source billions of years ago, and has spent the intervening time traveling through intergallactic space.) In any case, we don't even know exactly what our galaxy looks like now, because we are inside it!

We conclude this chapter with a little visual notation (Figure 2.10). Henceforth a flat two-dimen-

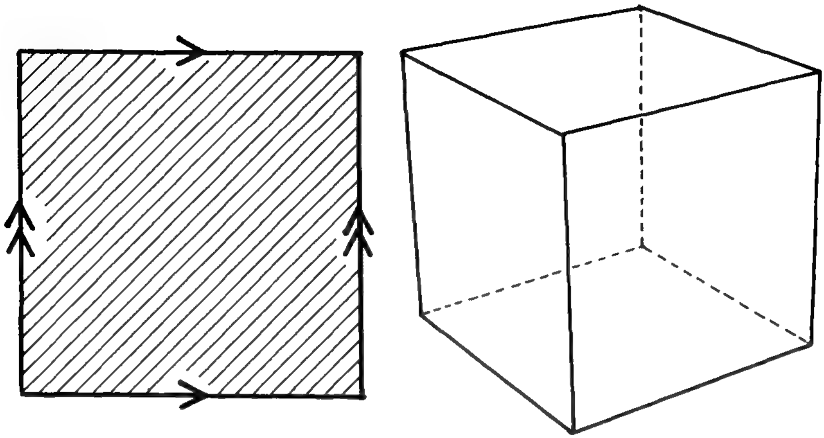


Figure 2.10 Representation of a flat two-dimensional torus (left) and a three-dimensional torus (right).

sional torus will be drawn as a square with arrows marked on its edges; you imagine the square's edges to be glued so that corresponding arrows match up. While it's possible to devise an analogous scheme for marking the faces of a cube, it isn't very practical. So we'll represent a three-torus simply by drawing a cube and stating that opposite faces are considered glued. By the way, a two-manifold like the two-dimensional torus is called a surface even though it isn't the surface *of* anything.

3

Vocabulary

This chapter explains five concepts basic to the study of manifolds:

1. Topology vs. geometry
2. Intrinsic vs. extrinsic properties
3. Local vs. global properties
4. Homogeneous vs. nonhomogeneous geometries
5. Closed vs. open manifolds

Don't worry about mastering these concepts right away! If you get the general idea now you can always refer back to this chapter later should the need arise.

Besides, later chapters will reinforce the ideas introduced here.

Most of the examples in this chapter will be surfaces, but the concepts apply to three-manifolds as well.

TOPOLOGY VS. GEOMETRY

Imagine a surface made of thin, easily stretchable rubber. Bend, stretch, twist, and deform this surface any way you like (just don't tear it). As you deform the surface, it will change in many ways, but some aspects of its nature will stay the same. For example, the surface at the far left in Figure 3.1, deformed as it is, is still recognizable as a sort of sphere,[2] whereas the surface to the far right is recognizable as a deformed two-holed doughnut. The aspect of a surface's nature that is unaffected by deformation is called the topology of the surface. Thus the two surfaces on the left in Figure 3.1 have the same topology, as do the two on the right. But the sphere and the two-holed doughnut surface have different topologies: no matter how you try you can never deform one to look like the other (remember—violence such as ripping one surface open and regluing it to resemble the other is *not* allowed).

A surface's geometry consists of those properties that *do* change when the surface is deformed. Cur-

---

[2]By a sphere we always mean just the surface, as opposed to a solid ball. Note that a sphere is intrinsically two-dimensional, while a solid ball is intrinsically three-dimensional.

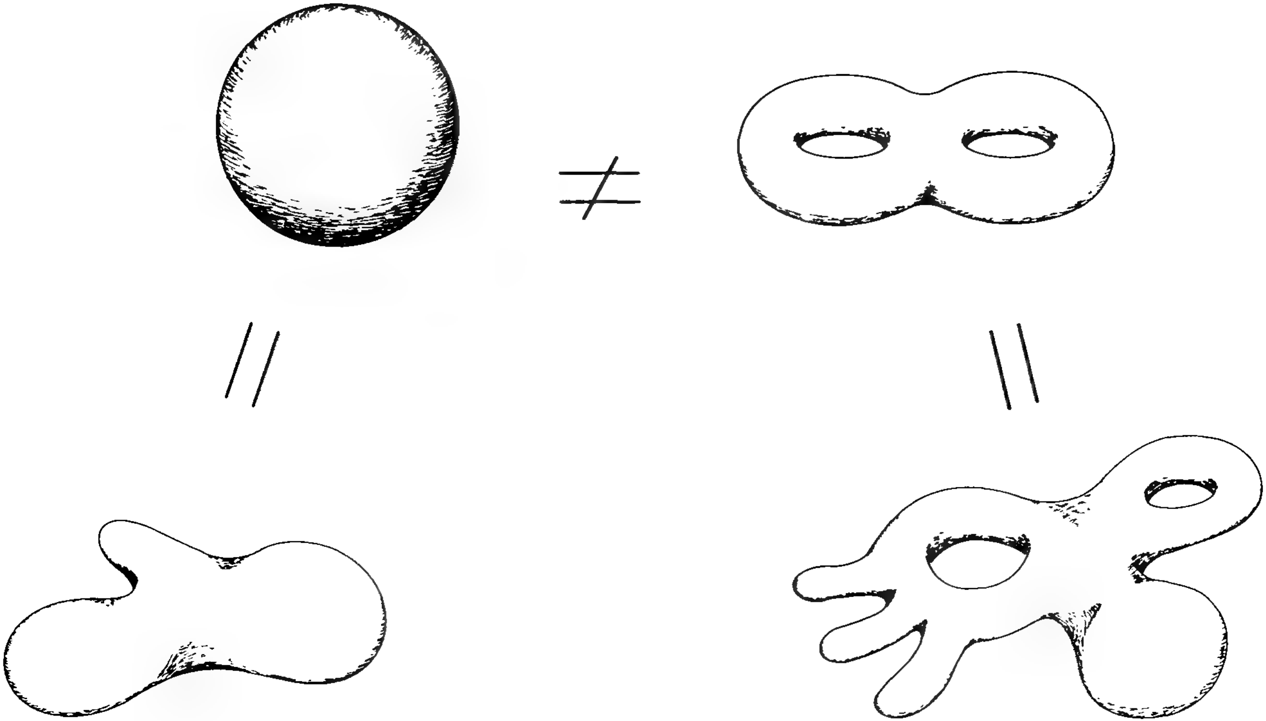


Figure 3.1 The two surfaces on the left are topologically indistinguishable, as are the two on the right.

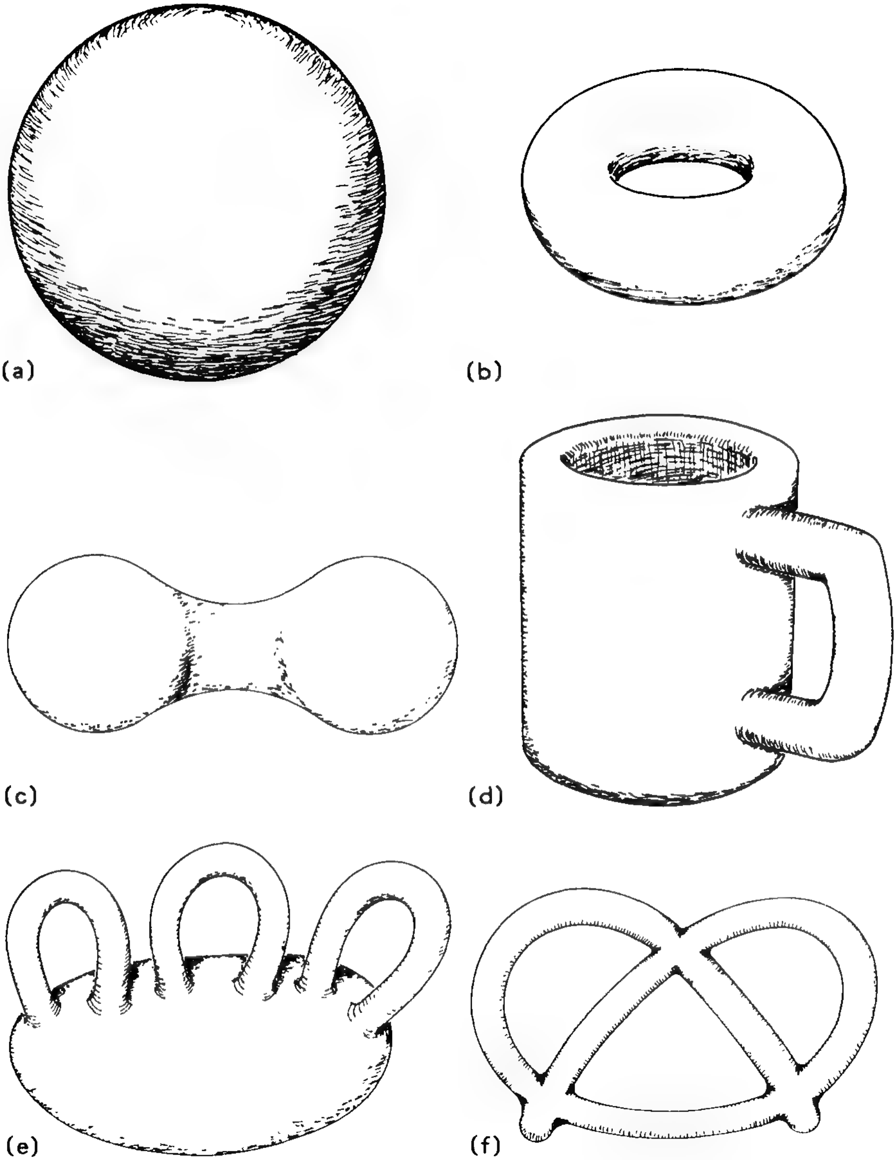
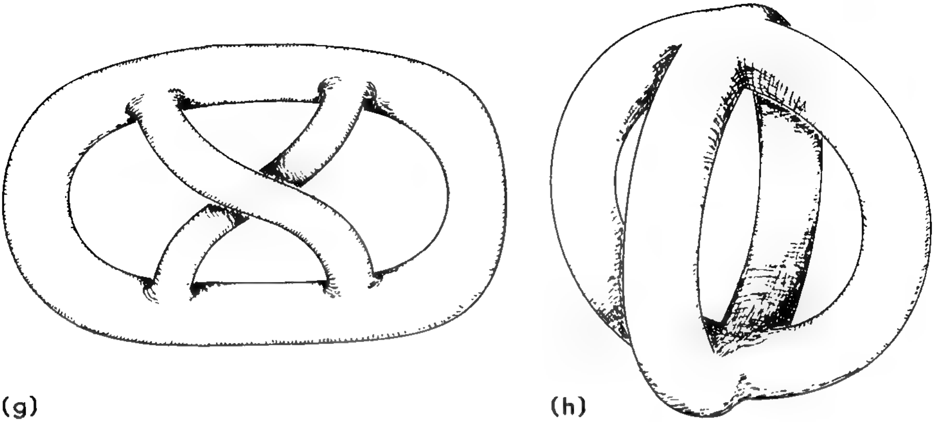


Figure 3.2 Which surfaces have the same topology?



(g) (h)

Figure 3.2 Continued.

vature is the most important geometrical property. Other geometrical properties include areas, distances and angles. An eggshell and a ping-pong ball have the same topology, but they have different geometries. (In this and subsequent examples the reader should idealize objects like eggshells and ping-pong balls as being truly two-dimensional, thus ignoring any thickness the real objects may possess.)

Exercise 3.1 Which of the surfaces in Figure 3.2 have the same topology? □

Exercise 3.2 In the story in Chapter 1, A Square discovered that in Flatland one can lay out two loops of thread that cross at only one point (namely downtown Flatsburgh). Did he discover a topological or a geometrical property of Flatland? □

If we *physically* glue the top edge of a square to its bottom edge, and its left edge to its right edge, then we will get a doughnut surface (Figure 3.3). The flat torus and the doughnut-surface torus have the same

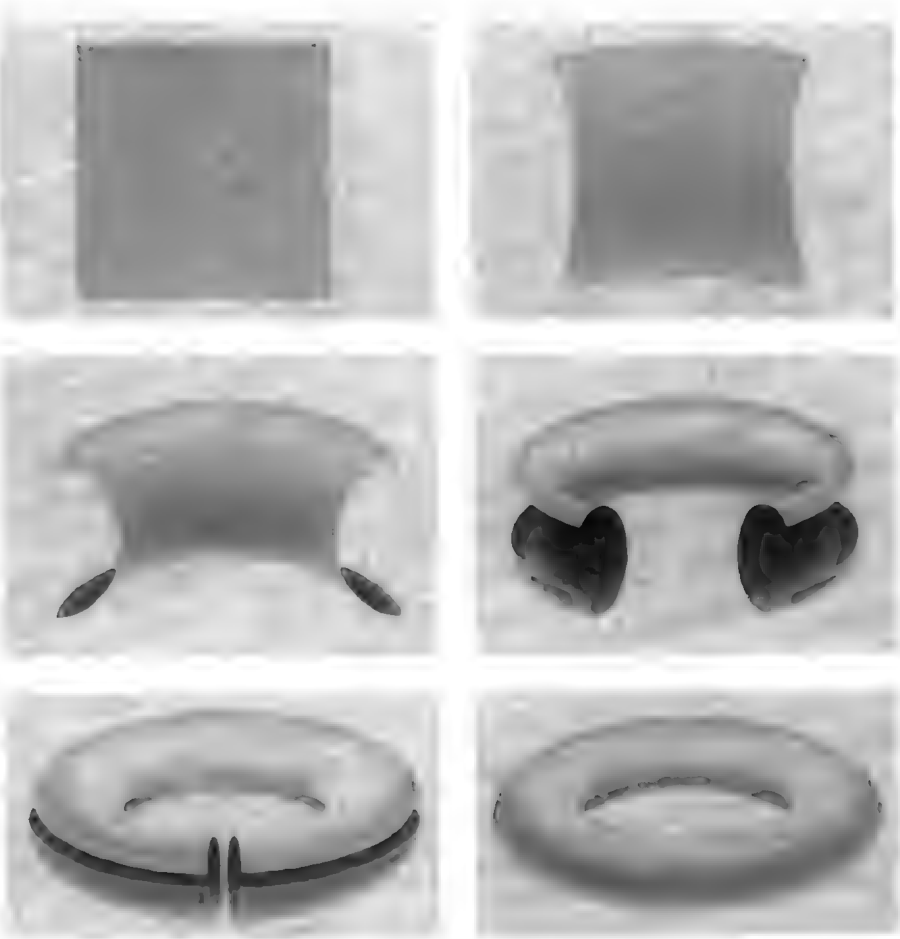


Figure 3.3 The flat torus and the doughnut surface have the same topology.

topology. They do not, however, have the same geometry. The doughnut surface is curved while the flat torus is obviously flat. In Figure 3.3 we had to deform the flat torus to get it to look like the doughnut surface.

The fact that the flat torus and the doughnut surface have the same topology explains why both are called tori ("tori," pronounced "tor-eye," is the plural of "torus"). It's only when we're interested in geometry that we distinguish the flat torus from the doughnut surface. In geometry the flat torus is vastly more important than the doughnut surface, so in this book, unless specified otherwise,

"torus" will mean "flat torus"

INTRINSIC VS. EXTRINSIC PROPERTIES

Figure 3.4 shows how to put a twist in a rubber band. The twisted band is topologically different from the original untwisted one.

At least from *our* viewpoint the bands are topologically different. In contrast, imagine how a Flatlander living in the band itself would see the cut-twist-and-reglue procedure. His whole world is the band; he has no idea that three-dimensional space exists at all. Thus he has no way to detect the twist. He sees the band get cut, and then—after a pause—get restored exactly to its original condition!

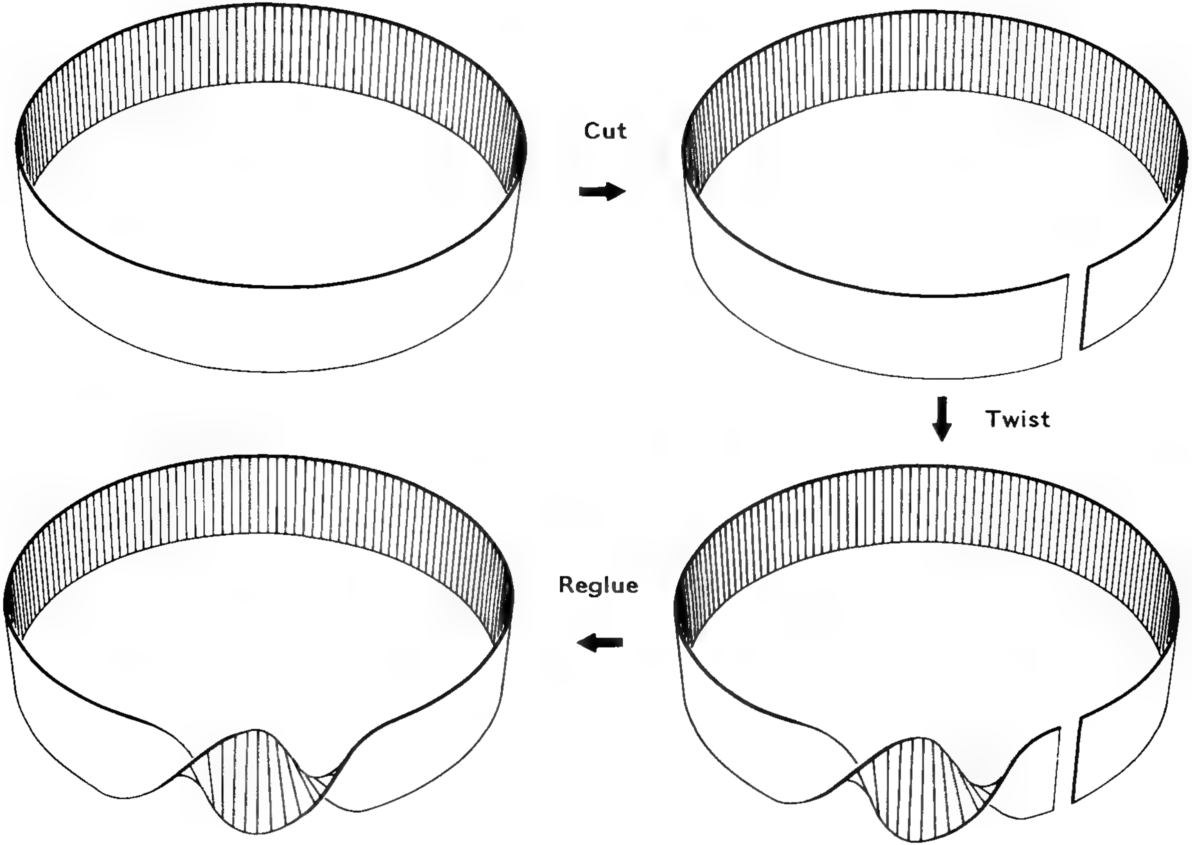


Figure 3.4 How to put a twist in a rubber band.

The *intrinsic topology* of the band has not changed, although its *extrinsic topology*—the way it's embedded in three-dimensional space—*has* changed.

In general, two surfaces have the same intrinsic topology if Flatlanders living in the surfaces cannot (topologically) tell one from the other. Two surfaces have the same extrinsic topology if one can be deformed *within three-dimensional space* to look like the other.

Exercise 3.3 All the surfaces in Figure 3.5 have the same intrinsic topology. Which have the same extrinsic topology as well? □

Exercise 3.4 Modify the cut-twist-and-reglue procedure of Figure 3.4 so that the intrinsic and extrinsic topology of the band both change. □

The intrinsic/extrinsic distinction also applies to the geometry of a surface. As an example, take a sheet of paper and bend it into a half-cylinder as shown in Figure 3.6. The *extrinsic geometry* of the paper has obviously changed. But the paper itself has not been deformed—its *intrinsic geometry* has not changed. In other words, a Flatlander living in a sheet of paper could not detect whether the paper was bent or not.

Here's an experiment to illustrate the above idea: Mark two points on a (flat) sheet of paper, and draw a straight line connecting them. The line represents the shortest path between the points on the flat paper. Now roll the paper into a cylinder (tape or glue it in

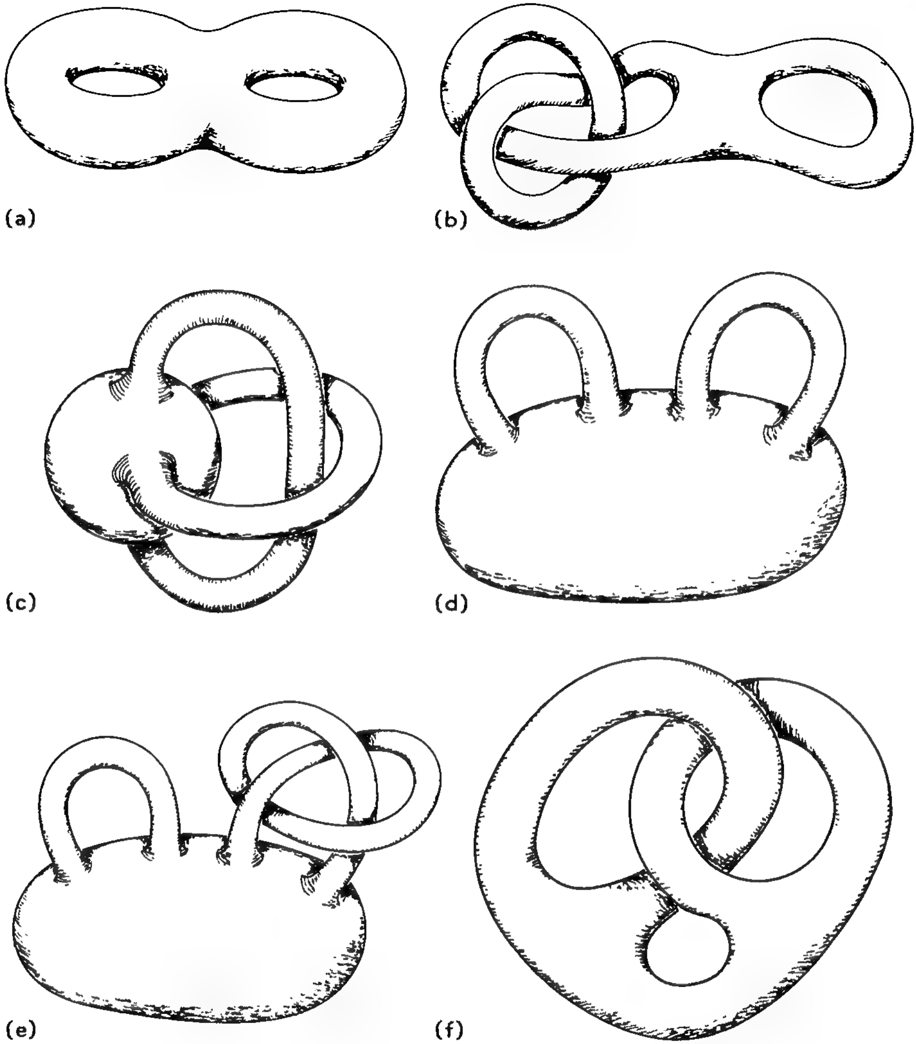


Figure 3.5 Which surfaces have the same extrinsic topology?

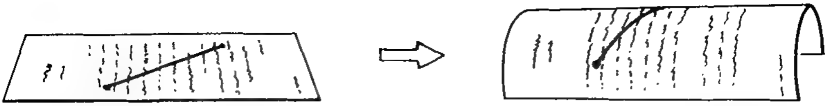


Figure 3.6 Bending a sheet of paper changes its extrinsic—but not its intrinsic—geometry.

position if necessary). Get a piece of thread, and wrap it around the cylinder from one point to the other. The thread will lie directly over the line you drew, indicating that the shortest path on the cylinder is the same as the shortest path on the flat paper. This is not surprising, because the "flat" paper and the "curved" cylinder both have the same intrinsic geometry.

Exercise 3.5 You can roll a piece of paper into a cylinder without deforming the paper. Can you also roll it into a cone without deformation? Can you wrap it onto a basketball without deformation? What does this tell you about the intrinsic geometries of the paper, the cylinder, the cone, and the basketball? □

Figure 3.7 shows three surfaces with different intrinsic geometries. A Flatlander could compare these surfaces by studying the properties of triangles drawn on them. (The sides of a triangle are required to be intrinsically straight in the sense that they bend neither to the left nor to the right. A Flatlander finds an (intrinsically) straight line in a surface the same way we Spacelanders do in our universe, e.g. by pulling taut a piece of thread, or by seeing how a beam of light

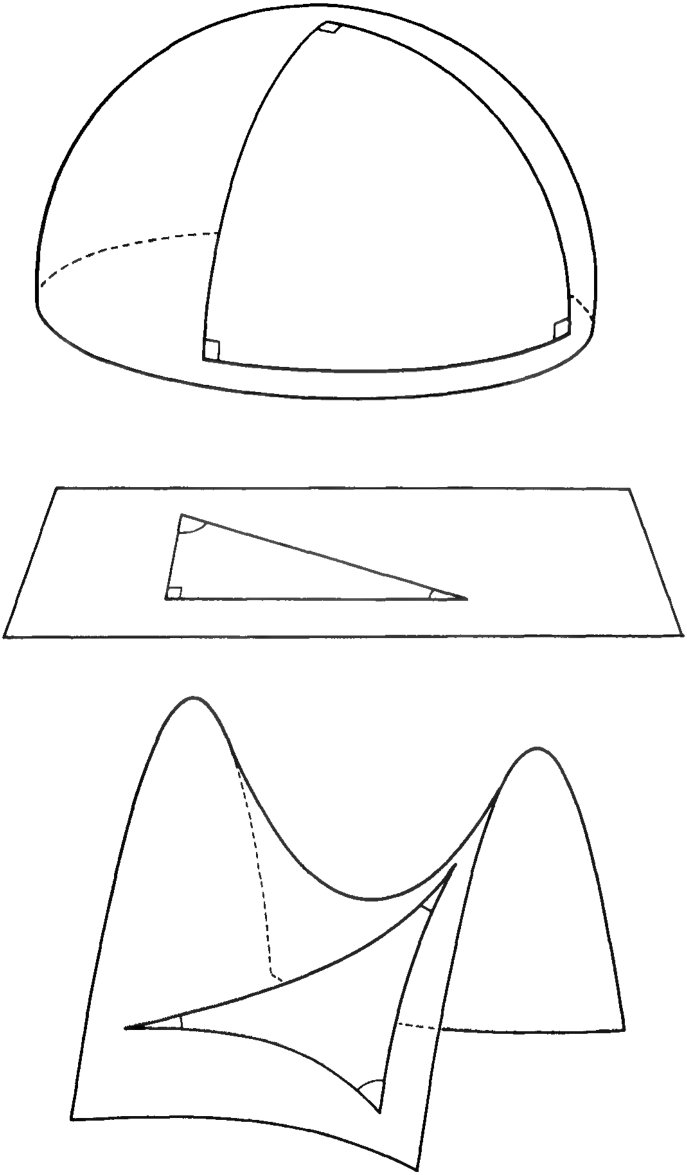


Figure 3.7 The hemisphere, the plane, and the saddle surface all have different intrinsic geometries.

travels. An intrinsically straight line is called a *geodesic*.) On the hemisphere, the sum of the angles of any triangle is greater than 180°. For example, the triangle shown in the figure has all its angles equal to 90°, so its angle-sum is 90° + 90° + 90° = 270°. In the plane, on the other hand, every triangle has angle-sum exactly equal to 180°. And in the saddle surface, all triangles have angle-sum less than 180°. Thus a Flatlander could experimentally determine which surface he lived in: he need only lay out a triangle and measure its angles! These properties of triangles are treated in detail in Chapters 9 and 10.

The mathematician Gauss carried out precisely this experiment in our own three-dimensional universe. (Later chapters will explain how a three-dimensional manifold can be curved. Gauss, though, was interested in the curvature of Earth's surface, and didn't expect to discover the curvature of space.) Gauss measured the angles in the triangle formed by the three mountain peaks Hohenhagen, Brocken, and Inselsberg. To within the accuracy of his measurements he found space to be intrinsically flat, i.e. the angles added to 180°. However, the universe is so vast and the Earth is so small that it would be impossible to detect any cosmic curvature on a terrestrial scale. Chapter 19 will treat the curvature of the universe in detail.

We Spacelanders can contemplate both the intrinsic and extrinsic properties of a surface. A Flatlander

does not have this option. His two-dimensional universe is all that's real and perceptible to him, so he naturally adopts an intrinsic viewpoint. For example, a Flatlander raised on a torus would have a very good intuitive understanding of what a torus is like, assuming of course that a trip across the torus was short enough to be an everyday sort of thing. He'd know intuitively that if he were going into the city, and then out to visit his aunt, and then home again, that it would be quicker to keep going in the same direction than to retrace his route.

Exercise 3.6 Imagine living in a three-torus universe where, after visiting friends in one galaxy, and doing a little exploration in another, it's quickest to keep on going to get home rather than turning back. □

We humans perceive our universe intrinsically, so when we study three-manifolds, such as the three-torus, we naturally visualize them intrinsically too. Because surfaces will guide us in our study of three-manifolds, it will also be useful to think of surfaces intrinsically. To this end we make the convention that *all surfaces will be studied intrinsically*, unless explicitly stated otherwise. Any extrinsic properties a surface may have will be ignored. (For example, we'll ignore the twist in the rubber band of Figure 3.4, and we'll ignore the bend in the sheet of paper of Figure 3.6.)

I'd like to insert a philosophical comment here. Even if the universe connects up with itself in funny

ways (if it's a three-torus, for example), this *doesn't* mean that it curves around in some four-dimensional space. The essence of the intrinsic point of view is that

> a manifold exists in and of itself, and needn't lie in any higher-dimensional space

LOCAL VS. GLOBAL PROPERTIES

A surface or three-manifold has both local and global properties. Local properties are those observable within a small region of the manifold, whereas global properties require consideration of the manifold as a whole. Try out this definition on the following exercise. Note that a sphere and a plane differ both locally and globally, and both topologically and geometrically.

Exercise 3.7 A society of Flatlanders lives on a sphere. They had always assumed they lived in a plane though, until one day somebody made one of the following discoveries. Which discoveries are local and which are global?

1. The angles of a triangle were carefully measured and found to be 61.2°, 31.7°, and 89.3°.
2. An explorer set out to the east and returned from the west, never deviating from a straight route.
3. As their civilization spread, the Flatlanders discovered the area of Flatland to be finite. □

The terms “local” and “global” are used most often in the phrases “local geometry” and “global topology.” For example, a flat torus and a doughnut surface have the same global topology, but different local geometries. A flat torus and a plane, on the other hand, have the same local geometry but different global topologies. A three-torus has the same local geometry as “ordinary” three-dimensional space, but its global topology is different. The Flatlanders in Chapter 1 were discovering various global topologies for Flatland, but when Gauss surveyed the mountain peaks he was investigating the local geometry of our universe in the region of the Earth.

We can use the local/global terminology to restate the definition of a manifold. A two-dimensional manifold (i.e. a surface) is a space with the local topology of a plane, and a three-dimensional manifold is a space with the local topology of “ordinary” three-dimensional space. All two-manifolds have the same local topology, and all three-manifolds have the same local topology, but the local topology of a two-manifold differs from that of a three-manifold. I should mention that three-manifolds serve as possible shapes for the universe precisely because their local topology matches that of ordinary space: we know almost nothing about the universe’s global topology or its local geometry, but it’s fair to assume that throughout the universe the local topology is just like the local topology of the “ordinary” space we occupy in the solar system.

Exercise 3.8 Define the concept of a one-dimensional manifold and give an example of one. □

Exercise 3.9 Compare an infinitely long cylinder to a plane. Do they have the same local geometry? (As usual we mean intrinsic local geometry, although you could also compare their extrinsic local geometries.) Do they have the same global topology? The same local topology? Which of the three types of discoveries listed in Exercise 3.7 could Flatlanders use to distinguish a cylinder from a plane? □

Exercise 3.10 Compare a three-torus made from a cubical room to one made from an oblong rectangular room. Do they have the same local topology? The same local geometry? The same global topology? The same global geometry? □

HOMOGENEOUS VS. NONHOMOGENEOUS GEOMETRIES

A *homogeneous* manifold is one whose local geometry is the same at all points. The local geometry of a nonhomogeneous manifold varies from point to point. A sphere is a homogeneous surface. The surface of an irregular blob is nonhomogeneous. A doughnut surface, while fairly symmetrical, is nonhomogeneous: it is convex around the outside but saddle-shaped near the hole. A flat torus, however, is homogeneous because it's flat at all points. The flat torus is more important in geometry than the doughnut surface precisely because it's homogeneous while the doughnut surface is not. Spheres are more important than sur-

faces of irregular blobs for the same reason. The sphere and the flat torus are the only homogeneous surfaces we have seen so far, but there will be plenty more. A major theme of this book is finding homogeneous geometries for manifolds that do not already have one.

Exercise 3.11 Is the three-torus a homogeneous three-manifold? □

CLOSED VS. OPEN MANIFOLDS

Intuitively, *closed* means finite and *open* means infinite. Try out your intuition on the following exercise. You can check your answers in the back of the book.

Exercise 3.12 Which of the following manifolds are closed and which are open?

1. A circle
2. A line (the whole thing, not just a segment)
3. A two-holed doughnut surface
4. A sphere
5. A plane
6. An infinitely long cylinder
7. A flat torus
8. Ordinary three-dimensional space
9. A three-torus □

Unfortunately there are two complications to the simple idea of closed and open. One is that anything with an edge, such as a disk, is technically not even a manifold (it's a so-called manifold-with-boundary) and

therefore does not count as either closed or open. Thus the terms "closed" and "open" imply that the manifold has no edges. This stipulation will not be an issue in this book, but in other books you may find that when an author makes a statement such as "the flat torus is a closed surface," he is emphasizing not that the flat torus is finite, but that it has no edges (unlike the square from which it was made, which does have edges).

The second complication is more interesting. It turns out that there are surfaces that are infinitely long, yet have only a finite area. A typical example is a doughnut surface with a so-called "cusp" (Figure 3.8). The cusp is an infinitely long tube that gets narrower as it goes. The first centimeter of cusp has a surface area of 1 square centimeter (cm^2), the next centimeter of cusp has an area of $\frac{1}{2}$ cm^2, the next an area of $\frac{1}{4}$ cm^2, and so on. Thus, the total surface area of the cusp is $1 + \frac{1}{2} + \frac{1}{4} + \cdots = 2$ cm^2. What's important is not that the area of the cusp is precisely

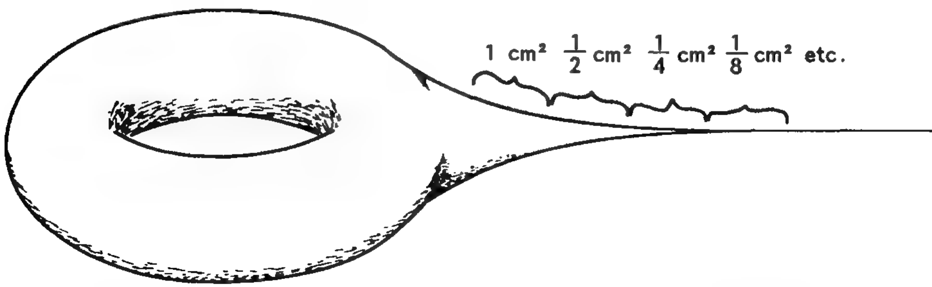


Figure 3.8 A doughnut surface with a "cusp" has a finite surface area even though the cusp is infinitely long.

2 cm^2, but that it's finite. For once you add in the (finite) area of the rest of the surface, you find that the surface as a whole has a finite area even though it's infinitely long. By convention a surface is classified as closed or open according to its distance across rather than its area, so the doughnut surface with a cusp is called open in spite of its finite area. After the following exercise we won't encounter any more cusps in this book.

Exercise 3.13 What would happen to A Square if he tried to take a trip down a cusp? □

This book deals mainly with closed manifolds, so from now on

"manifold" will mean "closed manifold"

unless explicitly stated otherwise.

4

Orientability

When our story left off in Chapter 1, our hero A Square and his fellow Flatlanders had just embarked on a Universal Survey of all of Flatland. The excitement was immense as the first survey party set out. But this excitement was nothing compared to the chaos that followed its return!

An old farmer living in an outlying agricultural district was the first one to run into the returning surveyors. He was going around a bend in the road and the surveyors were coming from the opposite direction. Fortunately no one was hurt in the collision. The

farmer was a little annoyed that the surveyors didn't have the courtesy to keep to the proper side of the road, but his anger was quickly overcome by his joy at seeing them back safely, and by his interest in hearing their tales of adventure. He accompanied them into town.

As they approached the Flatsburgh City Limits, one of the surveyors noticed that the "Welcome to Flatsburgh" sign, the one that announced the Rotary Club meetings and all, had been replaced by a backwards version of the same thing. A "hgrubstalF ot emocleW" sign, as it were.

"Those kids, always up to mischief," he chuckled.

"What's that you say?" asked the farmer, not sure he had heard properly.

"Oh, nothing. I was just amused at what some kids had done to that sign."

The farmer had no idea why the surveyor was so amused by the sign, but he decided not to make an issue of it.

The further the surveyors got into Flatsburgh, the more bewildered they became. *All* the signs were written backwards, and everyone, not just the old farmer, had taken to walking on the wrong side of the road. It was as if all of Flatsburgh had been mysteriously transformed into its mirror image while they were gone. Flatlanders in general tend to be superstitious, and for the surveyors this mirror reversal of Flatsburgh did not bode well.

Not that the citizens of Flatsburgh were any hap-

pier with the situation! They insisted that nothing the least bit unusual had happened in Flatsburgh while the surveyors were gone. It was the surveyors they found to be unusual, with their backwards writing and strange ideas about how Flatsburgh had been somehow transformed. In fact, they found the surveyors to be downright creepy, and, except for relatives and close friends, no one even wanted to go near them.

As you might suspect, the Universal Survey was called off, and for the next three years no one went more than shouting distance from the civilized parts of Flatland . . .

(*STORY TO BE CONCLUDED IN CHAPTER 5*)

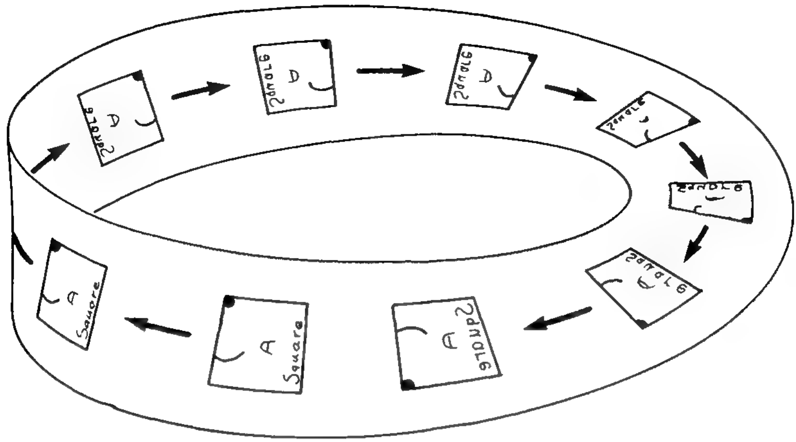
Exercise 4.1 Write a story in which you travel across the universe to an apparently distant galaxy, only to discover that you've made a complete trip around the universe and returned to our own galaxy. When you find Earth, you're startled to see that it looks like Figure 4.1. What do you see when you land? Describe a walk through your hometown. What do people think of you? □

In the Flatland story, each surveyor came back to Flatland as his own mirror image. To see just how this occurred, study Figure 4.2, which shows a swath of territory similar to the one the surveyors traversed. This swath of territory is a Möbius strip. [A true Möbius strip has zero thickness. If you mistakenly imagine it to have a slight thickness—like a Möbius strip

Figure 4.1 A mirror reversed Earth.

made from real paper—then you'll run into problems with A Square returning from his journey on the opposite side of the paper from which he started. As long as the Möbius strip is truly two-dimensional (i.e. no thickness) this problem does not arise.]

The question is, in what sort of surface could a Flatlander traverse a Möbius strip? A Klein bottle is one example. You can make a Klein bottle from a



Departure Return

Figure 4.2 When A Square travels around a Möbius strip he comes back as his mirror image.

square in almost the same way we made a flat torus from a square. Only now the edges are to be glued so that the arrows shown in Figure 4.3 match up. As with the flat torus, I don't mean that these gluings should actually be carried out in three-dimensional space; I mean only that a Flatlander heading out across one edge comes back from the opposite one. The top and bottom edges are glued exactly as in the flat torus: when a Flatlander crosses the top edge he comes back from the bottom edge and that's all there is to it. The left and right edges, though, are glued with a "flip." When a Flatlander crosses the left edge he comes back from the right edge, but he comes back mirror reversed. A Klein bottle contains many Möbius strips (see Figure 4.4).

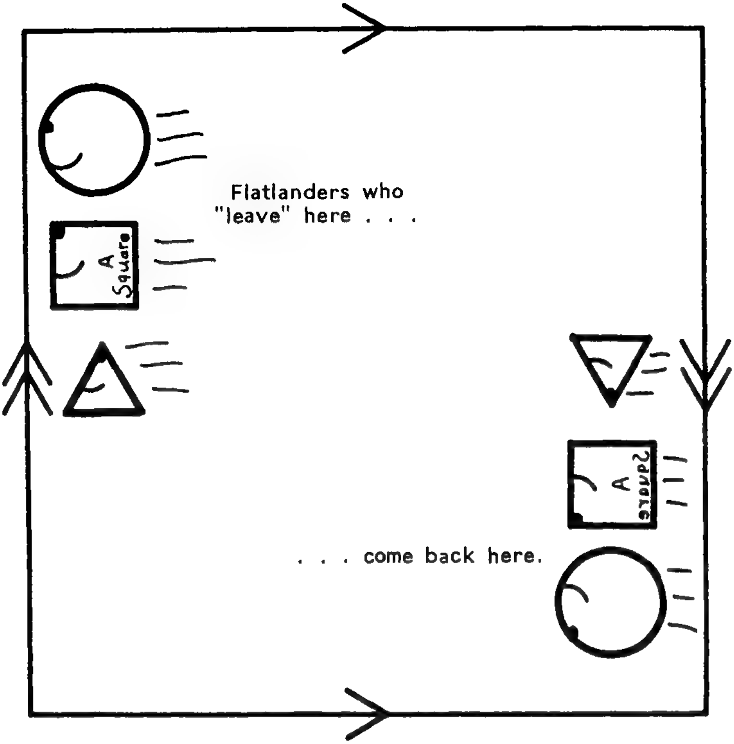


Figure 4.3 Glue the edges of this square so that the arrows match up and you'll get a Klein bottle. A Flatlander traveling off to the left comes back from the right as his mirror image.

Exercise 4.2 Which of the positions in Figure 4.5 constitute a winning three-in-a-row in Klein bottle tic-tac-toe? □

Exercise 4.3 Imagine the chessboard in Figure 2.6 to be glued to form a Klein bottle rather than a torus. Which black pieces does the white knight threaten now? Which black pieces threaten it? □

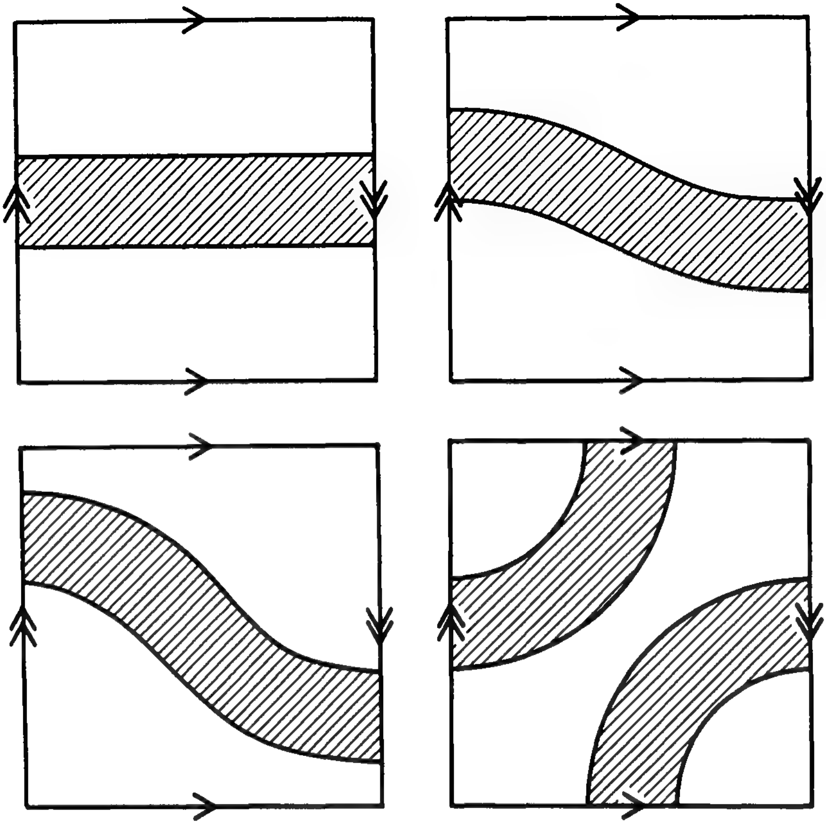


Figure 4.4 A Klein bottle contains many Möbius strips.

There's a nice way to analyze positions in Klein bottle tic-tac-toe and chess. For example, say there's a Klein bottle tic-tac-toe game in progress. The position is as shown on the left side of Figure 4.6 and it's X's turn to move. Rather than hastily taking the upper right hand square, X pauses to carefully analyze the situation. He notes that the board's top edge is glued to its bottom edge; therefore he draws a copy of the

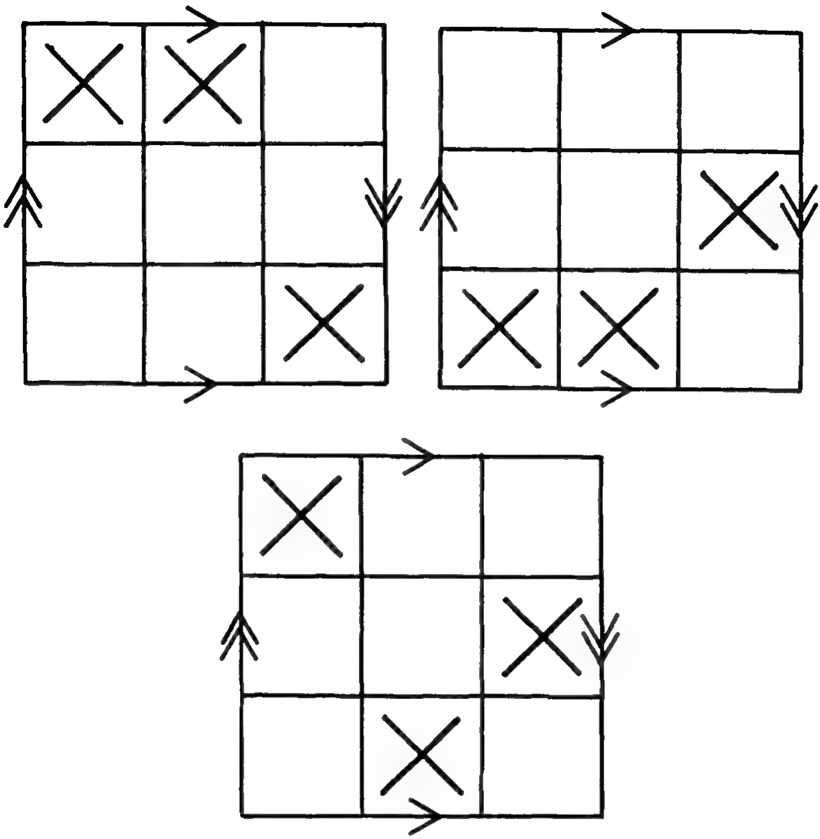


Figure 4.5 Which of these are winning positions in Klein bottle tic-tac-toe?

board above the original so that he can see more clearly how the top and bottom edges connect. He draws another copy below the original board for the same reason. He does the same on the left and the right, being careful to flip these copies top to bottom to account for the fact that he's playing tic-tac-toe on a Klein bottle and not on a torus. He continues this

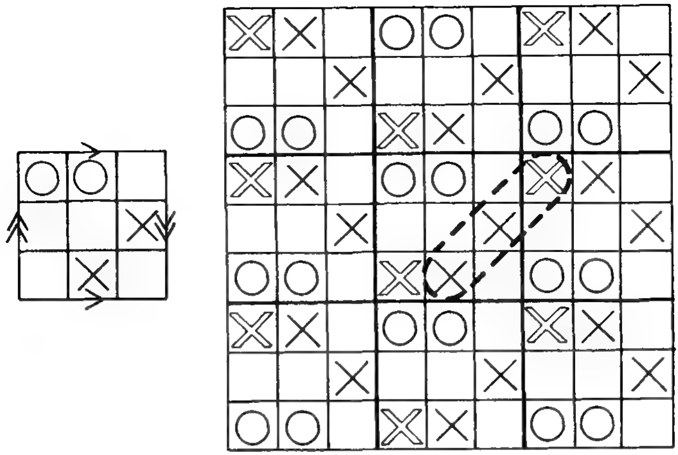


Figure 4.6 Analyzing a Klein bottle tic-tac-toe game.

process, attaching new copies of the board to the old ones. In principle he could continue forever, but he knows that nine copies are always enough to see clearly what is going on. After examining the final picture, X gleefully takes the lower left hand corner and wins immediately.

Exercise 4.4 Use the above technique to find X's best move in each situation shown in Figure 4.7. □

Exercise 4.5 Find a friend and play a few games of Klein bottle tic-tac-toe. □

Exercise 4.6 Find a friend and play a few games of Klein bottle chess. □

Exercise 4.7 When a bishop goes out the upper right hand corner of a Klein bottle chessboard, where does

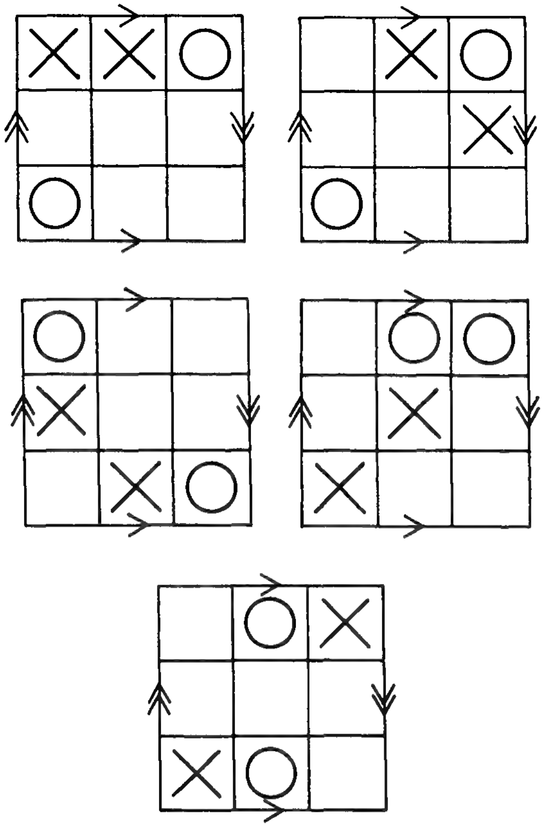


Figure 4.7 Find X's best move in each game.

he return? Hint: Label the corners of the board and draw a picture like Figure 4.6. □

Exercise 4.8 In Exercise 4.3 you found that the knight and one of the bishops threatened each other simultaneously on the Klein bottle chessboard. How can this be? Shouldn't a knight threaten only pieces on an oppositely colored square, while a bishop threat-

ens only pieces on the same color square? Can a knight and a bishop ever threaten each other simultaneously if the chessboard is constructed as in Figure 4.8? Can a knight and a rook simultaneously threaten each other on this new board? □

The global topology of most manifolds must be understood intrinsically. But we can assemble a Klein bottle in three-dimensional space in order to apprehend its global topology more directly. Imagine mak-

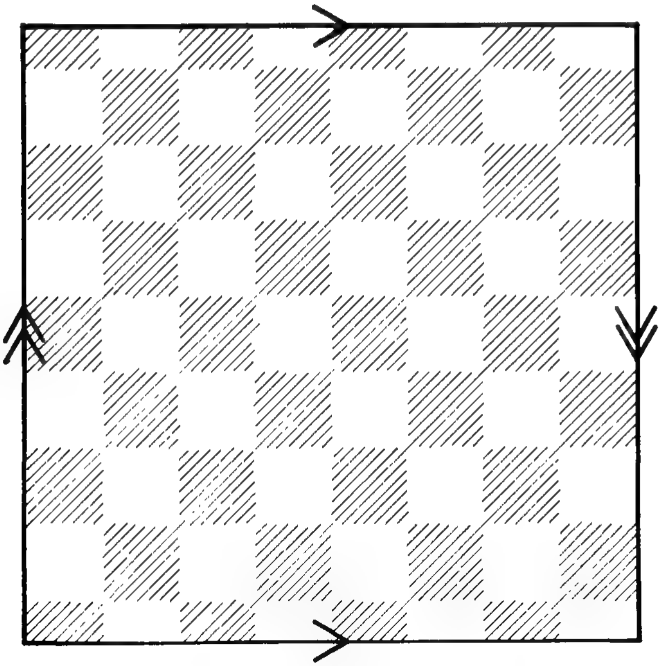


Figure 4.8 A new way to make a Klein bottle chessboard. The half-width row at the top gets glued to the half-width row at the bottom to become a normal row in the Klein bottle itself.

ing a Klein bottle from a square of rubber. Roll the square into a cylinder and glue the top edge to the bottom edge. That was the easy part. Now pass the cylinder through itself (as shown in Figure 4.9), and glue its ends together. The self-intersection is unpleasant, but there's no way to embed a Klein bottle in three-dimensional space without it. (As we'll see in Chapter 9, one *can* embed a Klein bottle in four-dimensional space with no self-intersection.) Note: Fig-

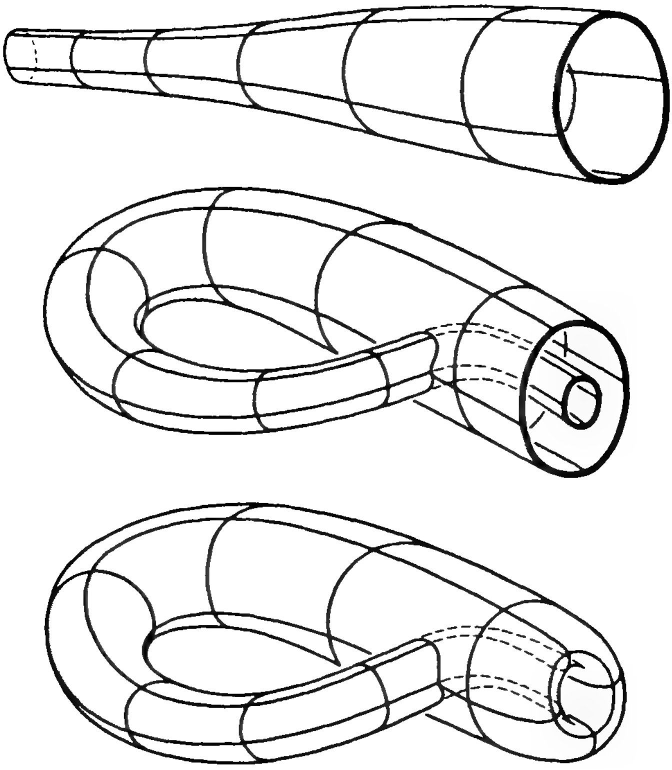


Figure 4.9 Gluing up a rubber Klein bottle.

ure 4.9 was included to make the Klein bottle's global topology a little more real, but for most purposes it's better to picture the Klein bottle as a square with opposite edges glued appropriately.

The local geometry of the Klein bottle is everywhere flat. Thus, a Flatlander doing local experiments could not distinguish a Klein bottle from either a torus or a plane. It's important to note that the Klein bottle is flat not only in the region corresponding to the middle of the square, but also in the regions where the edges and corners meet. Figure 4.10 shows how the four corners fit together. (By the way, the "seams" created by the gluing should all be erased. They are not part of the Klein bottle itself.) The Klein bottle is our

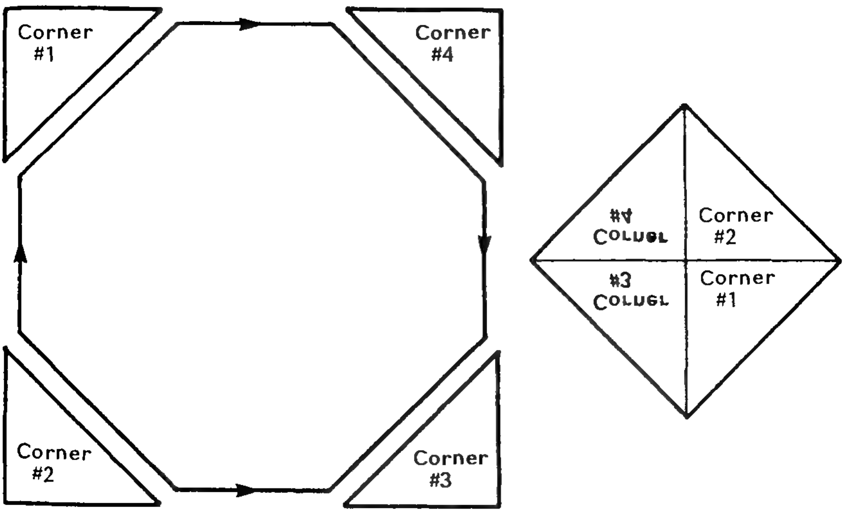


Figure 4.10 How the square's corners fit together in the Klein bottle.

third example of a homogeneous surface, the other two being the (flat) torus and the sphere.

A path in a surface or three-manifold which brings a traveler back to his starting point mirror-reversed is called an *orientation-reversing path*. Manifolds that don't contain orientation-reversing paths are called *orientable*; manifolds that do are called *nonorientable*. Thus, a sphere and a torus are orientable surfaces. A Klein bottle is a nonorientable surface. The three-torus is an orientable three-manifold. But what about a nonorientable three-manifold?

We can make a nonorientable three-manifold in much the same way that we made the Klein bottle. Start with the block of space inside a room. Imagine the left wall glued to the right wall, and the floor glued to the ceiling, just like in the three-torus. Only now imagine the front wall glued to the back wall with a side-to-side flip. If you walk through the front wall you'll return from the back wall mirror-reversed! Walk through it again and you'll come back in your usual condition.

Exercise 4.9 What do you see when you look through the back wall of this three-manifold? What about the other walls? □

Imagine this new three-manifold to contain a jungle gym like the one we built in the three-torus. This would be great fun to play in with a friend. Sometimes

when you ran into your friend he would be right-handed, and other times you'd find him to be left-handed. You could play a special form of tag in which catching someone doesn't count unless you can guess which hand is his left hand. Obviously there's a big advantage for people who part their hair in the middle, and T-shirts with writing on them would be out of the question.

Exercise 4.10 Think up other fun things to do in a nonorientable three-manifold. You could, for example, steal one shoe from your friend during the night, take it around an orientation-reversing path, and quietly replace it before dawn. □

The *projective plane* is a surface that is locally like a sphere, but has different global topology. It's made by gluing together the opposite points on the rim of a hemisphere (Figure 4.11). Figure 4.12 shows what this gluing looks like locally, along a short section of the rim. We can show the gluing along any section of the rim we like, but we can't show the entire gluing at once because of its peculiar global properties. Thus you should concentrate on understanding how opposite sections of rim fit together, rather than trying to visualize the whole thing at once the way you'd visualize a sphere. The most important thing is that the hemisphere's geometry matches up perfectly when opposite sections of rim are glued, so the projective plane has the same local geometry as a sphere, even along

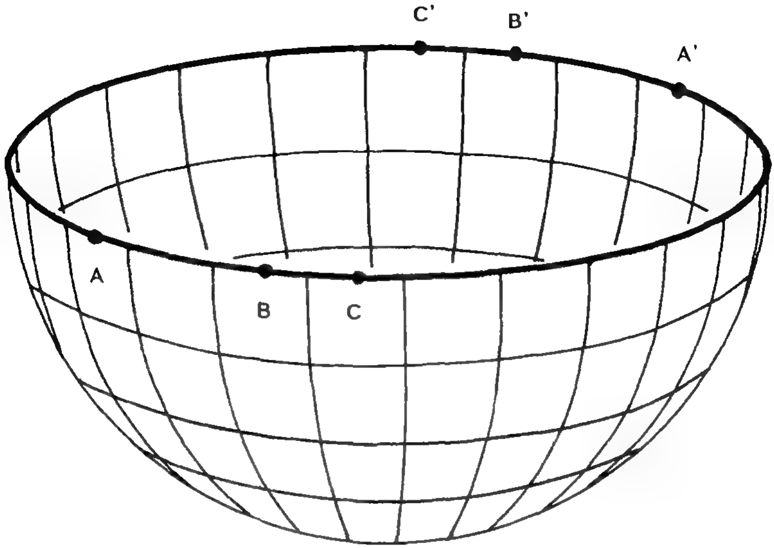


Figure 4.11 The projective plane is made by gluing together opposite points on the rim of a hemisphere.

the "seams" where the gluing took place. The projective plane is our fourth homogeneous surface.

Exercise 4.11 Is the projective plane orientable? That is, if a Flatlander crosses the "rim," does he come back normal or mirror-reversed? □

Exercise 4.12 A Flatlander lives on a projective plane, which we visualize as a hemisphere with opposite rim points glued. The Flatlander's house is at the "south pole." One day he leaves his house and travels in a straight line (i.e., a geodesic) until he gets back home again. At what point along the route is he furthest from home? ⧉

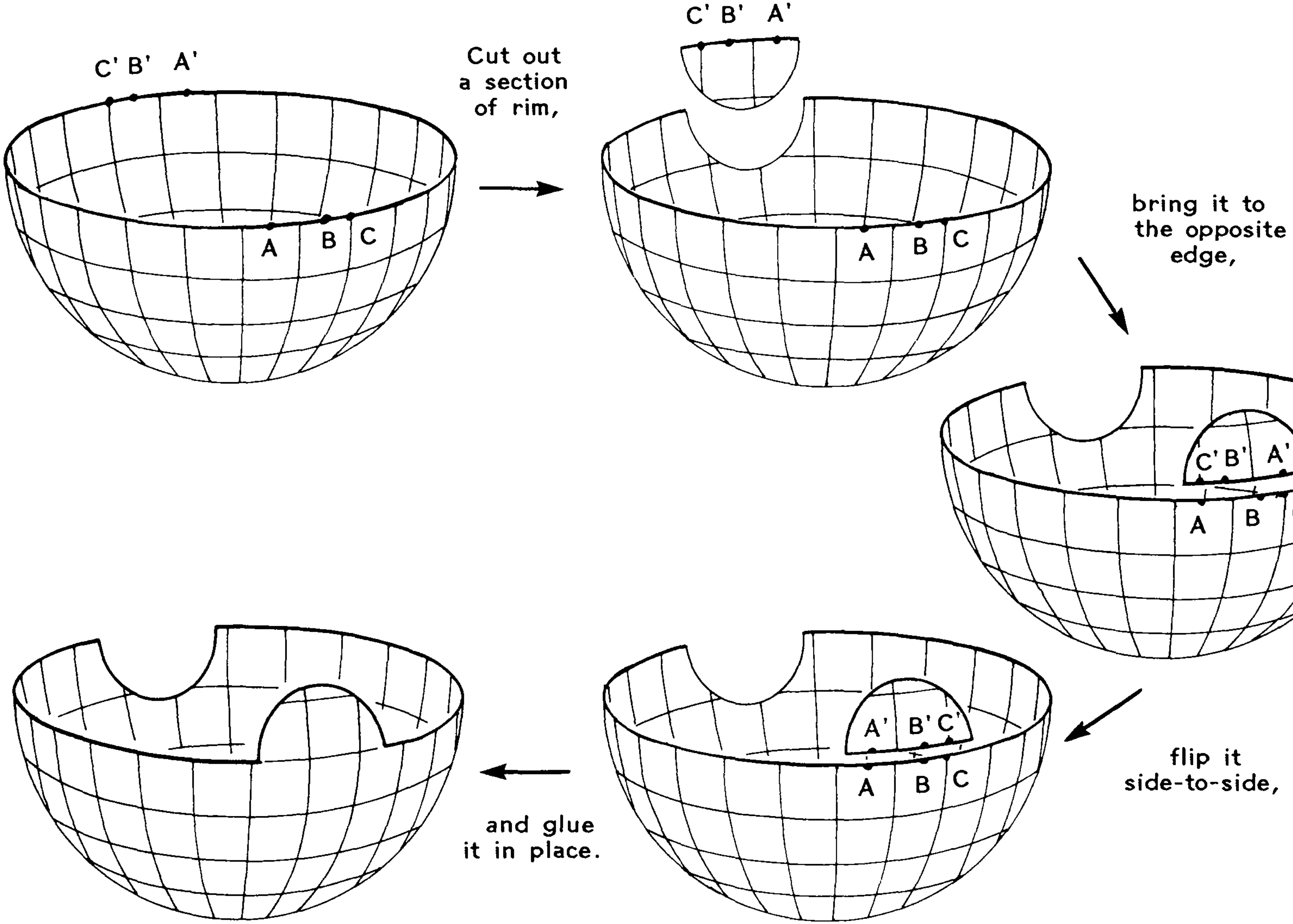


Figure 4.12 How to glue opposite sections of rim.

Exercise 4.13 A society of Flatlanders lives on a projective plane. They plan to build two fire stations. For maximal effectiveness the fire stations should be as far apart as possible. Where might the Flatlanders build them? (Be careful: opposite points on the rim of a hemisphere represent the same point in the projective plane.) How should three fire stations be positioned for maximal effectiveness? □

Exercise 4.14 A Flatlander knows he lives on either a sphere or a projective plane. How can he tell which it is? A second Flatlander knows he lives on either a projective plane or a Klein bottle; how can he decide? □

Exercise 4.15 So far we have seen four homogeneous surfaces: the sphere, the torus, the Klein bottle, and the projective plane. Use them to fill in the table in Figure 4.13 □

If we are interested in only the topological properties of the projective plane, we can flatten the hemisphere into a disk, still remembering to glue opposite boundary points (Figure 4.14). The main advantage of doing this is that a disk is easier to draw than a hemisphere.

Still working topologically we can construct *projective three-space* by gluing opposite boundary points of a solid three-dimensional ball. We'll study its ge-

	orientable	nonorientable
curved local geometry		
flat local geometry		

Figure 4.13 The sphere, torus, Klein bottle and projective plane can fill this table. Which goes where?

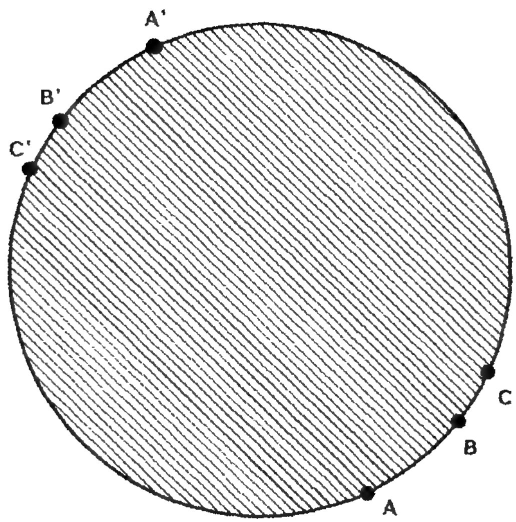


Figure 4.14 Topologically, a projective plane is a disk with opposite boundary points glued.

ometry in Chapter 14 when we study the geometry of the hypersphere.

Exercise 4.16 Is projective three-space orientable? If you cross the boundary, how do you come back? □

Exercise 4.17 Is orientability a local or a global property? Is it topological or geometrical? □

5

Connected Sums

Conclusion of the Flatland story:

The mirror-reversed surveyors adapted to their new condition more quickly than most had expected. The hardest part was learning to write properly, but even this became routine after a while. And with their increased competence came a greater acceptance on the part of the community. Things returned to normal.

In fact, as the years went by people even got a little adventurous. Almost every week somebody or another was heading out on an expedition. There were, of course, occasional incidents of explorers com-

ing back mirror-reversed, but this was no longer a disaster. The reversed explorers were quickly rehabilitated. Besides, the reversal incidents were limited to those who passed through a certain "Reversing Region." The rest of Flatland seemed harmless enough, and trips there became quite common.

To protect travelers from accidental reversal, the Reversing Region was marked with clumps of stones spaced ten paces apart along its boundary. Once this was done, even the most timid Flatlanders enjoyed traveling about in the safe regions.

It wasn't long before an official survey of the safe regions was undertaken. The surveyors found that the safe regions resembled a doughnut surface (Figure 5.1), in accordance with one of the earliest proposed theories on the shape of Flatland.

Others were quick to point out, however, that this didn't mean that Flatland as a whole was a doughnut surface. For example, if the Reversing Region also re-

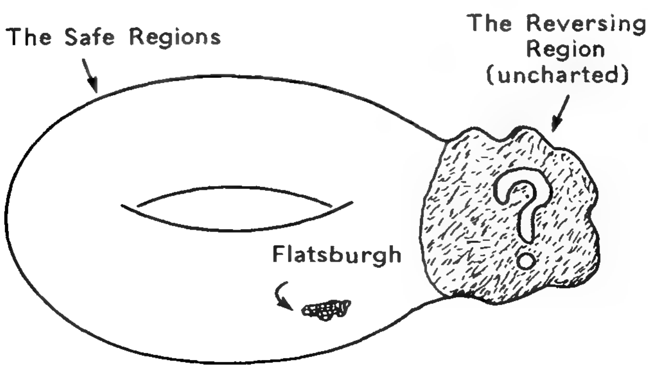


Figure 5.1 The safe regions of Flatland were charted first.

sembled a doughnut surface, then Flatland would be a two-holed doughnut surface (Figure 5.2).

Curiosity overcame fear, and a survey of the Reversing Region was begun. Only the boldest of the surveyors volunteered for the job. It wasn't that they were afraid of getting reversed—that was common enough by now. They were afraid of getting reversed-*twice*! The popular consensus was that a second reversal would result in certain death. (There was a minority opinion that a second reversal would simply restore the victim to his original state, but this opinion didn't sell as well in the newspapers.)

The survey was divided into two stages. The purpose of the first stage was to get a rough idea of just how big the Reversing Region was, and to divide it

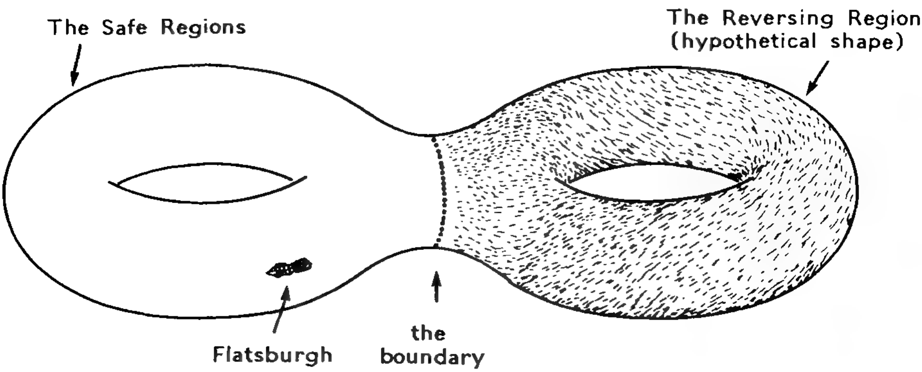


Figure 5.2 People were quick to point out that if the safe region and the Reversing Region each resembled a doughnut surface, then Flatland as a whole would be a two-holed doughnut surface.

into sectors to be mapped in detail during the second stage.

The first stage went smoothly, even though three of the surveyors came back mirror-reversed. But these reversed surveyors were brave enough to go back into the Reversing Region to help with the detailed surveying of the second stage. In fact, they even drew lots to see who was the bravest and would go back in first!

After the completion of the first stage the Reversing Region was divided into eight sectors. During the second stage a separate team was sent to each sector to map it in detail. The whole operation had an air of Russian roulette, with each team wondering whether they were the ones mapping the dangerous sector that did the reversing. To everyone's surprise—and relief—all eight teams reported their respective sectors to be perfectly normal!

It was only when they compiled, consolidated, and compared the data from the different sectors that things got mysterious. They found the sectors connected up as shown in Figure 5.3. The mysterious thing was that Sector 1 connected to Sector 8, not Sector 7; it was Sector 2 that connected to Sector 7! They connected in such a confusing way!

Eventually confusion gave way to enlightenment. The Flatlanders realized that the mirror-reversal phenomenon wasn't so mysterious after all. It was simply that the space of Flatland connected up with itself in such a way that anyone taking a trip around the Reversing Region would come back with his left side

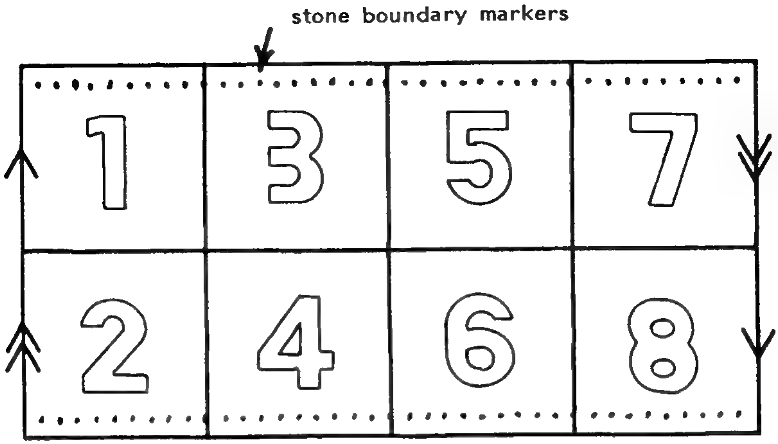


Figure 5.3 How the eight sectors pieced together.

where his right side was, and his right side where his left side was. The Flatlanders had discovered the Möbius strip!

This was an immense intellectual achievement. But it was a very practical achievement as well. The reversed surveyors were all sent on a trip around the Reversing Region to restore them to their original condition. Thereafter the Reversing Region was used mainly for pranks and other amusements.

Thus, the Universal Survey was complete: the surveyors had established beyond a doubt that Flatland consists of two regions, one a Möbius strip, and the other resembling a torus. The Flatlanders lived happily and peacefully forever after.

THE END

> NOTE: All surfaces in this chapter will be considered topologically, so you may bend and twist them however you like!

As the Flatlanders pointed out in Figure 5.2, a two-holed doughnut surface bears a strong resemblance to two one-holed doughnut surfaces stuck together. In fact, we can make a two-holed doughnut surface from two one-holed ones by cutting a disk out of each and gluing together the exposed edges (see Figure 5.4). This operation is called a connected sum.

Exercise 5.1 What do you get when you form the connected sum of a two-holed doughnut surface and a one-holed doughnut surface? What is the connected sum of a six-holed doughnut surface and an eleven-holed one? □

Exercise 5.2 What do you get when you form the connected sum of a two-holed doughnut surface and a sphere? How about a Klein bottle and a sphere? A projective plane and a sphere? □

Exercise 5.3 The purpose of this exercise is to find out what you get when you cut a disk out of a projective plane; in Exercise 5.4 you will use this information to deduce what the connected sum of two projective planes is. Get some scratch paper, and work your way through the following steps, drawing a picture for each one.

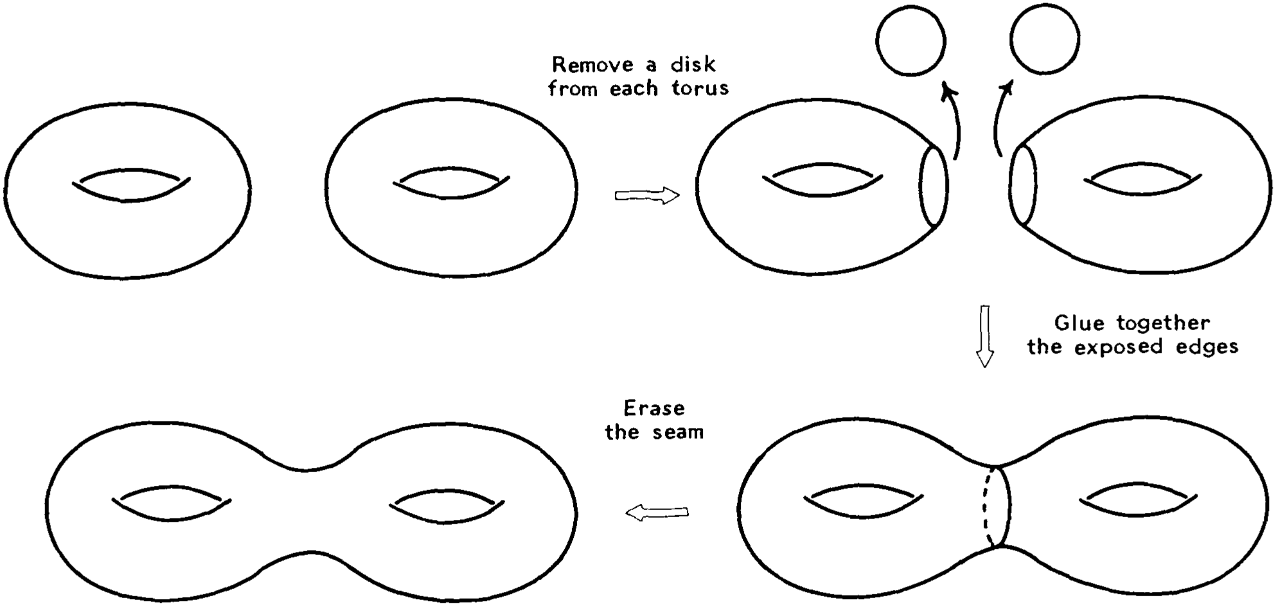


Figure 5.4 How to convert two one-holed doughnut surfaces into a single two-holed one.

1. Draw a (topological) projective plane as a disk with opposite edge points glued.
2. Remove a small disk from the center of the projective plane.
3. Cut what remains into two curved pieces as shown in Figure 5.5. Label the edges with arrows as shown.
4. Straighten each curved piece into a rectangle. We're studying the *topological* properties of the projective plane, so it's OK to bend the pieces as if they were made of rubber. Just don't lose track of which arrows are on which edges.
5. Physically glue together the two long edges labeled with single arrows. You'll have to flip one piece over to do this.

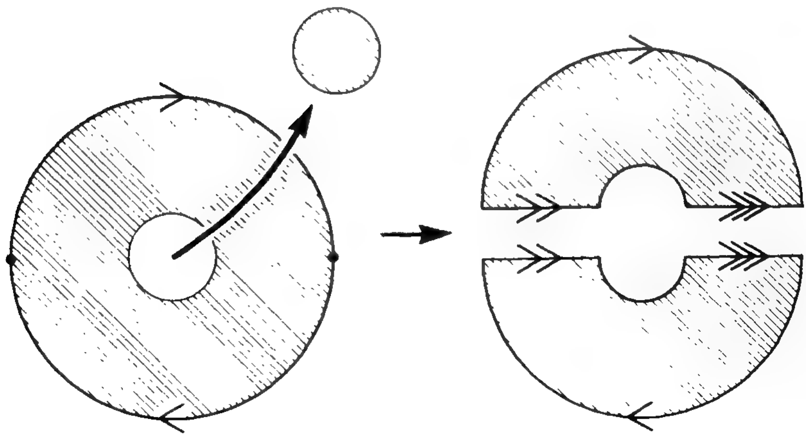


Figure 5.5 Remove a disk from a projective plane. Then cut what remains into two pieces as shown.

6. Physically glue together the edges with the double and triple arrows. You'll have to do this in three-dimensional space—it can't be done in the plane.

The thing you end up with is topologically identical to the projective plane with the disk removed. What is its usual name? ◻

Exercise 5.4 What is the connected sum of two projective planes? Hint: You start with two projective planes, cut a disk out of each, and get two of the things you discovered in the previous exercise. Now refer to Figure 5.6 and the following limerick to help decide what you get when you glue the two things together.

A mathematician named Klein
Thought the Möbius strip was divine.
Said he, "If you glue
The edges of two
You'll get a weird bottle like mine." ◻

Exercise 5.5 In the story at the beginning of this chapter, Flatland was the connected sum of what two surfaces? ◻

Most simple manifolds have shorthand names, usually written with a superscript to indicate their dimension. For example, the surfaces we've studied are

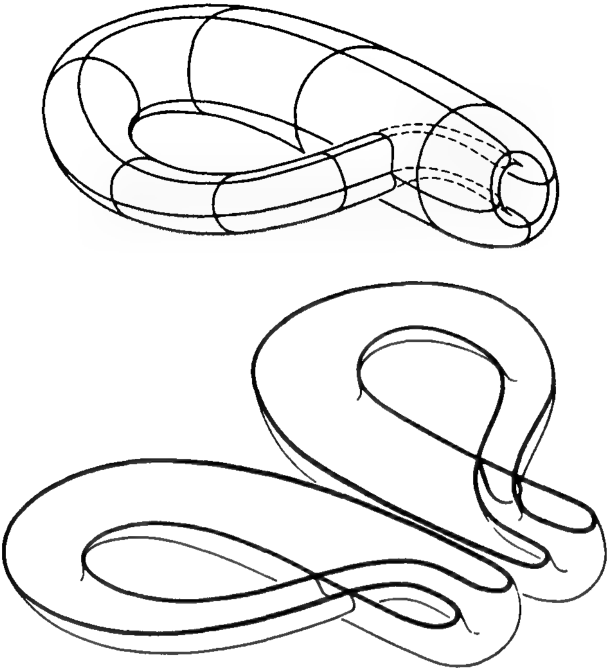


Figure 5.6 Cutting a Klein bottle in two.

E^2	The (Euclidean) plane
S^2	The sphere
T^2	The torus
K^2	The Klein bottle
P^2	The projective plane
D^2	The disk

The abbreviations are pronounced “E-two,” “S-two,” “T-two,” etc. By the way, there is no topological difference between a doughnut surface and a flat torus, so the abbreviation “T^2” may refer to either or both of them depending on the context. The three-manifolds we’ve seen are

E^3	"Ordinary" three-dimensional (Euclidean) space
T^3	The three-torus
D^3	A solid ball (i.e. a three-dimensional "disk")
P^3	Projective three-space

The nonorientable three-manifold we studied has a name, too; its name describes its structure as a so-called "product" and will be revealed in the next chapter. Even one-dimensional manifolds have abbreviations, namely

E^1	The line
S^1	The circle
I	The interval, i.e. a line segment with both endpoints included

The connected sum operation is abbreviated by a "#" symbol. For example, a two-holed doughnut surface is $T^2 \# T^2$ because it's topologically the connected sum of two tori. ($T^2 \# T^2$ is read "the connected sum of two tori" or simply "T-two connect-sum T-two".) Similarly, a three-holed doughnut surface is $T^2 \# T^2 \# T^2$. The topology of Flatland is succinctly written as $T^2 \# P^2$. One can even write equations with this notation. For example, in Exercise 5.4 you found that $P^2 \# P^2 = K^2$.

Exercise 5.6 State the results of Exercise 5.2 in the above notation. □

Sometime in the 1860s, mathematicians discovered that *every* conceivable surface is a connected sum of tori and/or projective planes! (The sphere counts as

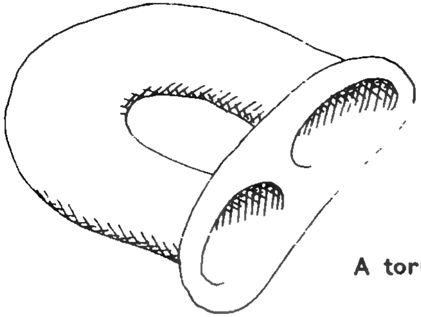
a connected sum of zero tori and zero projective planes. I know this sounds hokey, but it is convenient. And it's not that unreasonable in light of Exercise 5.2.) In other words, the table below provides a complete list of all possible surfaces. Any surface you might come up with is topologically equivalent to a surface in the table. The Klein bottle, for example, is equivalent to $P^2 \# P^2$, which occurs as the third entry in the first row.

		Number of Projective Planes				
		0	1	2	3	$\cdots$
Number of Tori	0	S^2	P^2	$P^2 \# P^2$	$P^2 \# P^2 \# P^2$	$\cdots$
	1	T^2	$T^2 \# P^2$	$T^2 \# P^2 \# P^2$	$T^2 \# P^2 \# P^2 \# P^2$	$\cdots$
	2	$T^2 \# T^2$	$T^2 \# T^2 \# P^2$	$T^2 \# T^2 \# P^2 \# P^2$	$\vdots$	
	3	$T^2 \# T^2 \# T^2$	$T^2 \# T^2 \# T^2 \# P^2$	$\vdots$		
	$\vdots$	$\vdots$	$\vdots$			

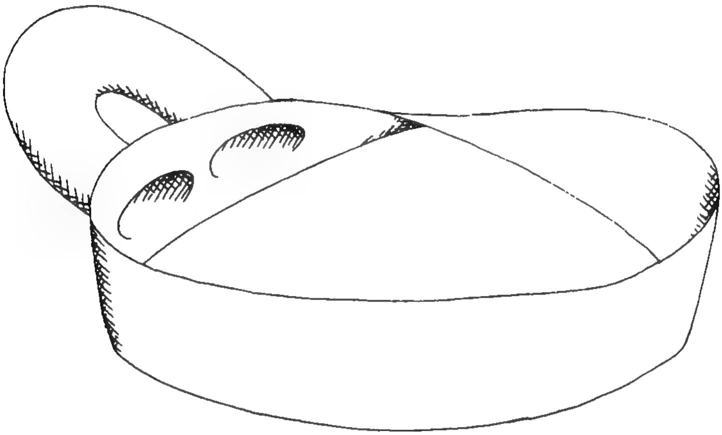
Exercise 5.7 Find a surface in the table that is topologically equivalent to (a) $K^2 \# P^2$, (b) $K^2 \# T^2$, (c) $K^2 \# K^2$. □

Does this list contain duplications? For example, might $T^2 \# P^2$ really be the same surface as, say, $P^2 \# P^2 \# P^2$? Surprisingly enough, the list does contain duplications, and its true that $T^2 \# P^2 = P^2 \# P^2 \# P^2$!

Exercise 5.8 In this exercise you discover that $T^2 \# P^2 = K^2 \# P^2$, and in Exercise 5.9 you'll use this information to deduce that $T^2 \# P^2 = P^2 \# P^2 \# P^2$. Convince yourself that each picture in Figure 5.7 is what its

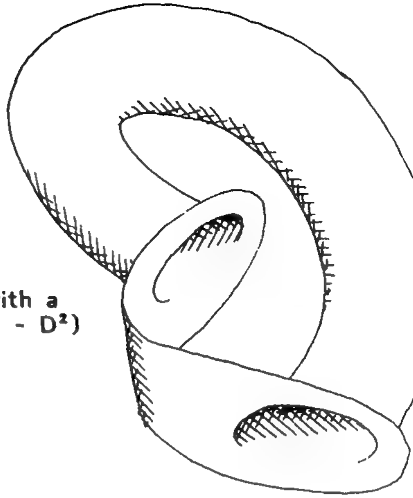


A torus with a disk removed
(T^2 - D^2)

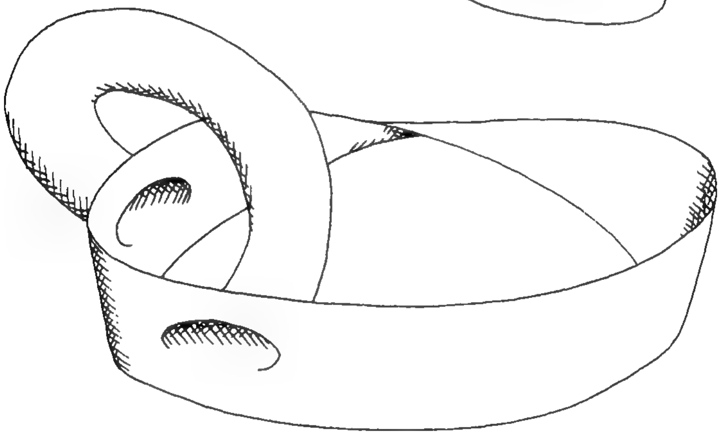


The connected sum of a torus
and a Möbius strip
(T^2 # Möbius)

Figure 5.7 Check that each picture is what its caption says it is.



A Klein bottle with a disk removed $(K^2 - D^2)$



The connected sum of a Klein bottle and a Möbius strip (K^2 # Möbius)

Figure 5.7 Continued.

caption says it is. From Exercise 5.3 you can deduce that gluing a disk to the edge of a Möbius strip converts the Möbius strip into a projective plane. Therefore T^2 # Möbius will become $T^2 \# P^2$ if you glue a disk to its edge, and K^2 # Möbius will become $K^2 \# P^2$ if you glue a disk to its edge. Since T^2 # Möbius and K^2 # Möbius are topologically the same (study Figure 5.8

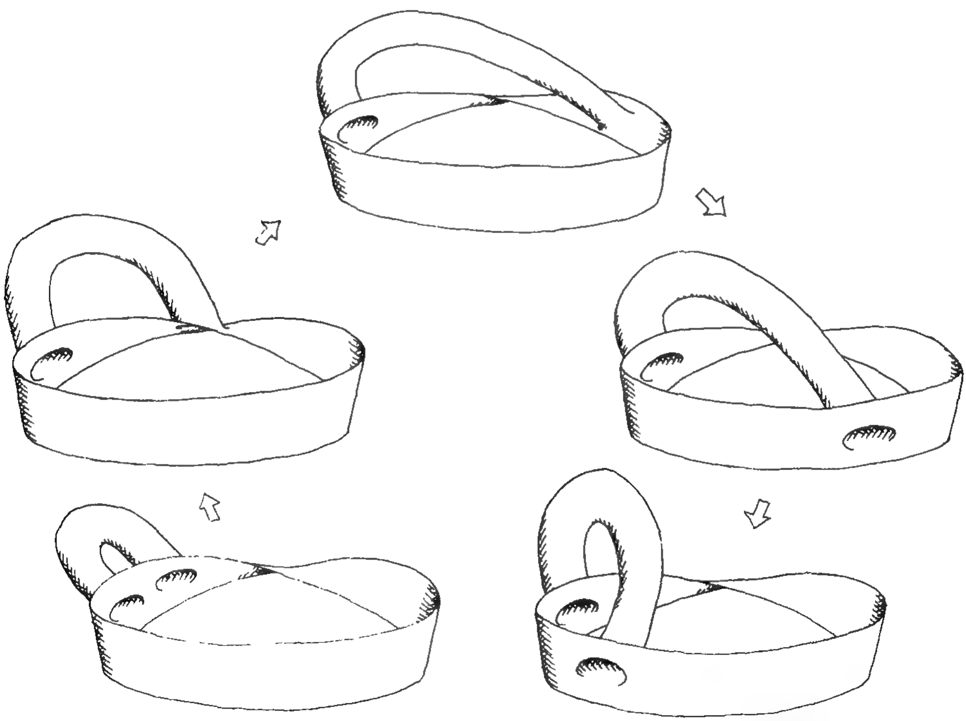


Figure 5.8 T^2 # Möbius and K^2 # Möbius can be deformed one into the other. Therefore they are topologically the same. *Real* rubber would of course break if you tried this sort of deformation, but in topology you needn't worry about such complications.

to see why), $T^2 \# P^2$ and $K^2 \# P^2$ must also be the same. (Oddly enough, this does *not* mean that T^2 and K^2 are the same! Unlike for addition and multiplication of numbers, there is no "cancelation law" for connected sums.) □

Exercise 5.9 Verify that $T^2 \# P^2 = P^2 \# P^2 \# P^2$ as claimed a couple paragraphs back. (Hint: This is easy! Simply combine Exercises 5.4 and 5.8.) □

Exercise 5.10 Assuming every surface is a connected sum of tori and/or projective planes, deduce that every surface is a connected sum of either tori *only* or projective planes *only*. That is, every surface is topologically equivalent to some surface on the following two column list:

$$
\begin{array}{cc}
\multicolumn{2}{c}{S^2} \\
T^2 & P^2 \\
T^2 \# T^2 & P^2 \# P^2 \\
T^2 \# T^2 \# T^2 & P^2 \# P^2 \# P^2 \\
T^2 \# T^2 \# T^2 \# T^2 & P^2 \# P^2 \# P^2 \# P^2 \\
\text{etc.} & \text{etc.}
\end{array}
$$

Where does $T^2 \# P^2$ (the surface in the Flatland story) appear on this list? Where is $T^2 \# K^2$? $P^2 \# S^2$? $S^2 \# S^2$? Which surfaces on the list are orientable?

In Chapter 12 we'll see that all the surfaces on this new list really are different. □

Exercise 5.11 Match each surface in Column A to a topologically equivalent surface in Column B.

Column A	Column B
$T^2 \# S^2$	$P^2 \# P^2$
K^2	$K^2 \# P^2$
$S^2 \# S^2 \# S^2$	$S^2 \# S^2$
$P^2 \# T^2$	$P^2 \# P^2 \# P^2 \# K^2$
$K^2 \# T^2 \# P^2$	T^2

□

One can also talk about a connected sum of two three-manifolds (you remove a solid ball from each and glue the remaining three-manifolds together along the exposed spherical boundary).

This Page Intentionally Left Blank

6

Products

A cylinder is the product of a circle and an interval. It qualifies as such because (see Figure 6.1) it is both

1. A bunch of intervals arranged in a circle, and
2. A bunch of circles arranged (in this case stacked) in an interval

More concisely, a cylinder is the product of a circle and an interval because it is both (1) a circle of intervals and (2) an interval of circles. It is abbreviated as

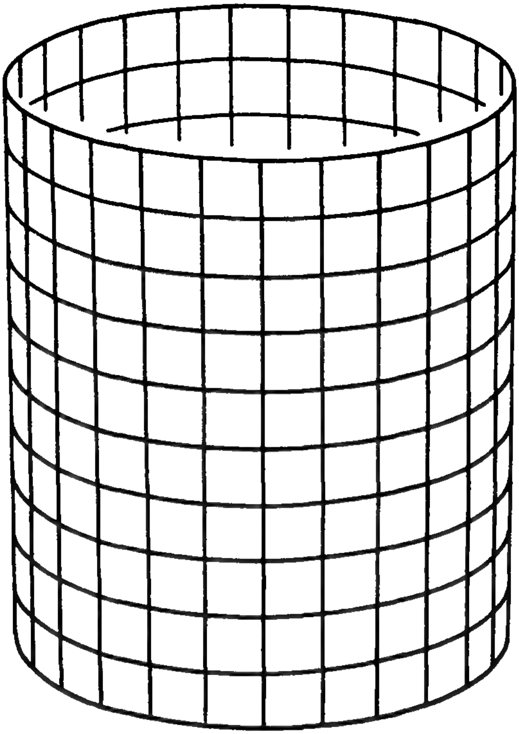


Figure 6.1 A cylinder is the product of a circle and an interval. (Recall that an "interval" is a line segment with both endpoints included.)

$S^1 \times I$ (pronounced "a circle cross an interval" or simply "S-one cross eye").

A torus is a second example of a product. It's the product of one circle (drawn dark in Figure 6.2) with another (drawn light). This is because the torus is a circle of circles in two different ways: both as a dark circle of light circles and as a light circle of dark circles. This fact is abbreviated by the equation $T^2 =$

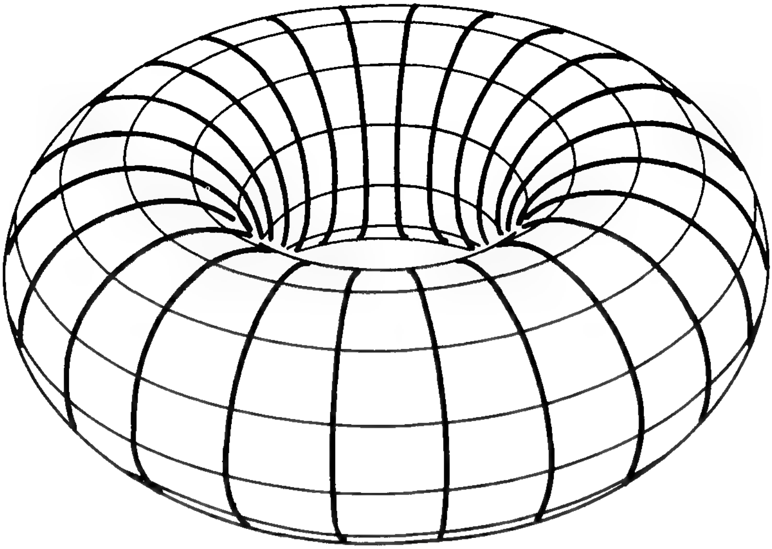


Figure 6.2 A torus is a circle cross a circle.

$S^1 \times S^1$ (read "a torus is a circle cross a circle" or simply "T-two equals S-one cross S-one"). The torus in Figure 6.2 is drawn as a doughnut surface, which is OK since we are interested only in its topology.

[By the way, the torus is the only closed surface (with no edges) that is a product. The reason is that a two-dimensional product must be the product of two one-dimensional things, and the circle is the only one-dimensional thing available that has neither endpoints (like an interval) nor infinite length (like a line); thus $S^1 \times S^1$ is the only two-dimensional product having neither an edge nor an infinite area.]

Exercise 6.1 What are the usual names for each of the following products:

1. $I \times I$
2. $E^1 \times E^1$
3. $S^1 \times E^1$
4. $E^1 \times I$. □

Exercise 6.2 Is the Möbius strip a product?

Exercise 6.3 What's $D^2 \times S^1$? Work topologically. □

There is a connection between products of manifolds and products of numbers in the sense of multiplication. For example, 15 is the product of 3 and 5, and the 15 sheep in Figure 6.3 form both 3 rows of 5 and 5 columns of 3.

Our first example—the cylinder—is a product not only in the topological sense but in the geometrical sense as well. To be precise, it qualifies as a geometrical product because it satisfies the following three conditions:

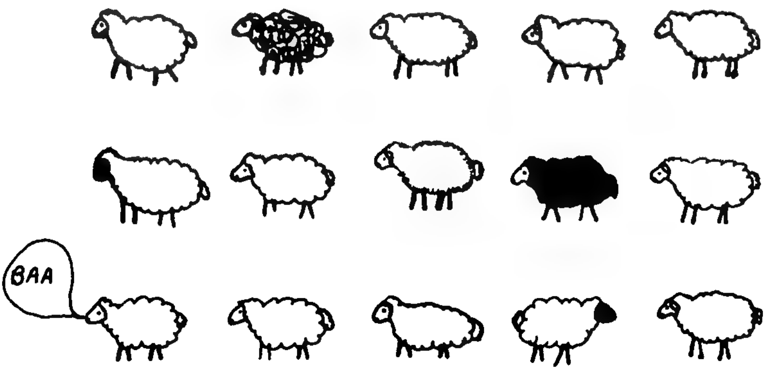


Figure 6.3 Fifteen is the product of three and five.

1. All the circles are the same size.
2. All the intervals are the same size.
3. Each circle is perpendicular to each interval.

Exercise 6.4 Figure 6.4 shows three renditions of $S^1 \times I$ that are topological—but not geometrical—products. Which of the above three conditions is violated in each case? □

Exercise 6.5 Draw one version of $I \times I$ that is a geometrical product and another version that is only a topological product. □

The doughnut surface in Figure 6.2 is not a geometrical product because the light circles are not all the same size. In fact, we can't draw a geometrical $S^1 \times S^1$ in three-dimensional space at all. The best we can do is to draw a cylinder as in Figure 6.1 and imagine that the top is glued to the bottom. The gluing converts each vertical interval into a circle, thereby converting $S^1 \times I$ into $S^1 \times S^1$. We know this new product is a geometrical one because

1. All the original circles are the same size.
2. All the intervals that get converted into circles are the same size.
3. Each circle is perpendicular to each interval.

But we also know that this cylinder with its ends glued together is a flat torus. Therefore a geometrical $S^1 \times S^1$ is a flat torus!

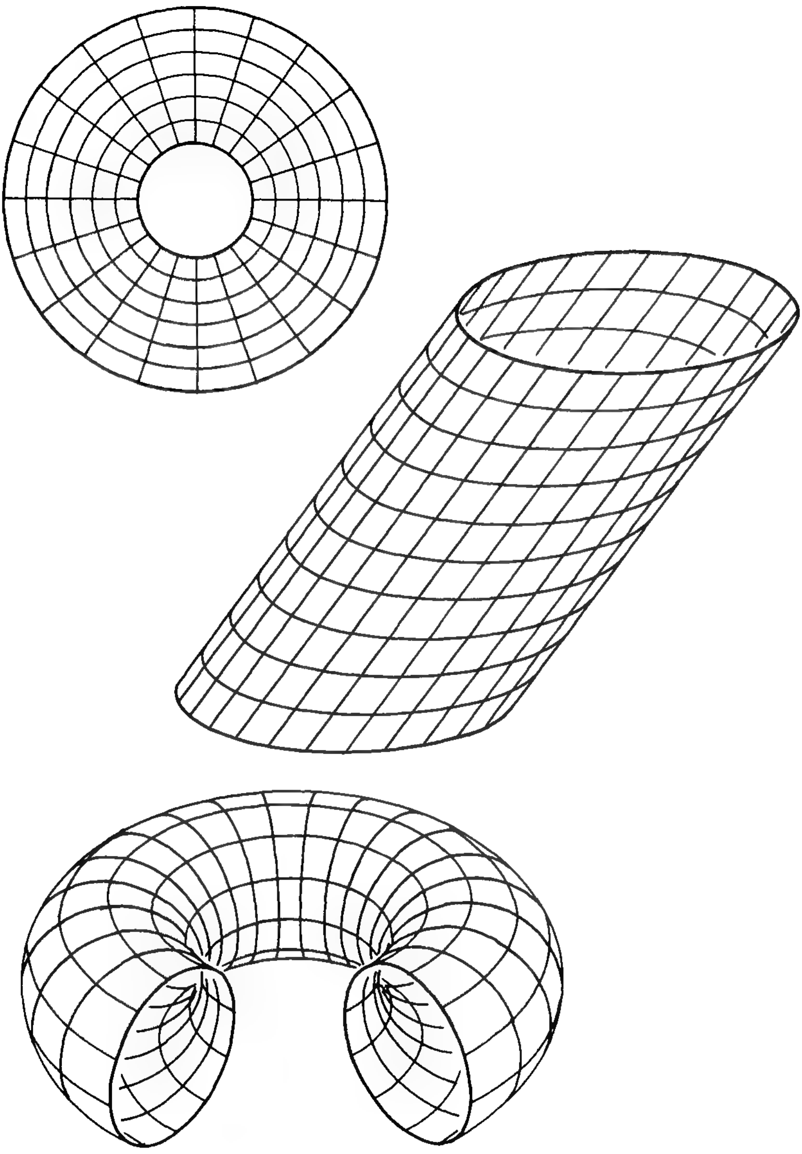


Figure 6.4 Three versions of $S^1 \times I$ that are topological—but not geometrical—products.

Our main reason for studying products is to better understand three-manifolds. It's true that the vast majority of three-manifolds *aren't* products, but many of the simplest and most interesting ones are. For example, the three-torus is the product of a two-torus and a circle (in symbols $T^3 = T^2 \times S^1$). Here's how to see it. Recall that a three-torus is a cube with opposite faces glued. Imagine this cube to consist of a stack of horizontal layers as shown in Figure 6.5. When the cube's sides get glued, each horizontal layer gets con-

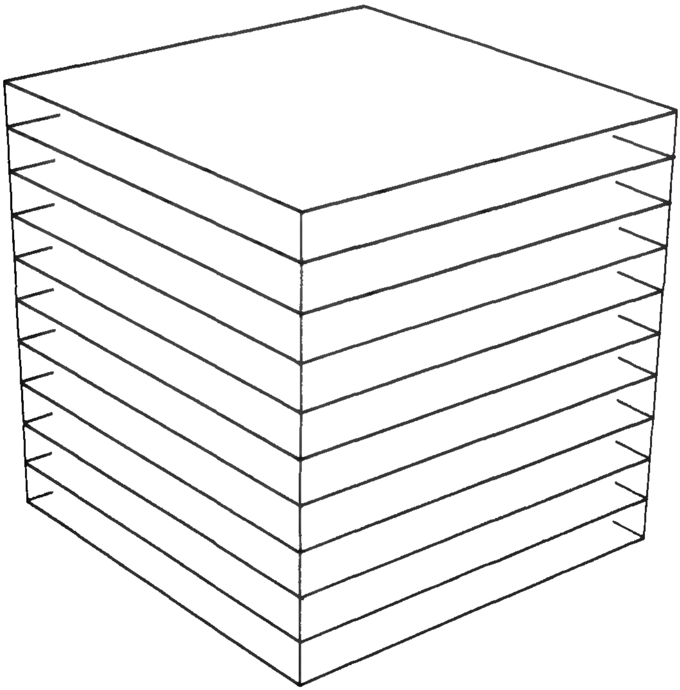


Figure 6.5 A three-torus is the product of a two-torus and a circle.

verted into a torus—a flat torus in fact. At this stage we have a stack of flat tori. When the cube's top is glued to its bottom, this stack of tori is converted into a circle of tori. We still have to check that T^3 is a torus of circles as well. To do this, imagine the cube to be filled not with horizontal layers, but with vertical intervals, like lots of spaghetti standing on end (Figure 6.6). When the cube's sides get glued this square of intervals becomes a torus of intervals, and when the top and bottom are glued it becomes a torus of circles

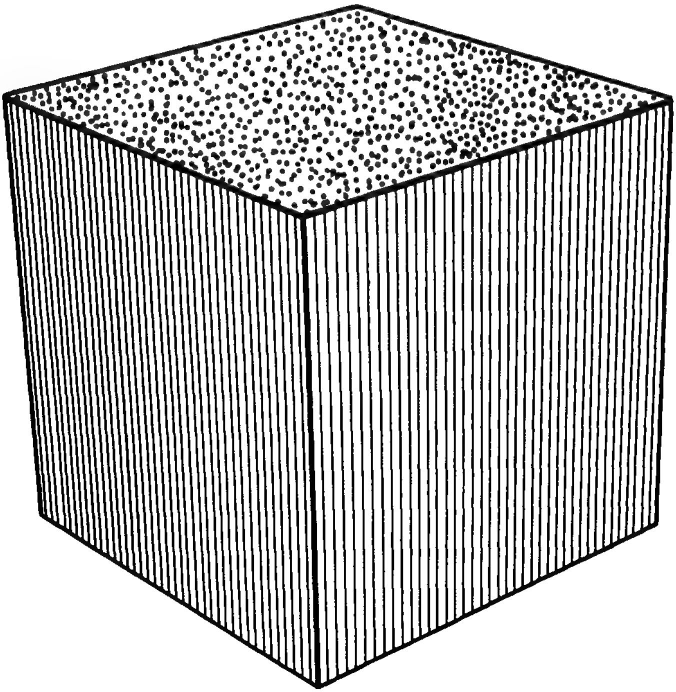


Figure 6.6 To see that T^3 is a torus of circles, first imagine the cube to be filled with spaghetti standing on end.

as required. Thus the three-torus is both a circle of tori and a torus of circles, so it is the product of a torus and a circle.

The three-torus is, in fact, a geometrical product because

1. All the horizontal tori are the same size (Figure 6.5).
2. All the vertical circles are the same size (Figure 6.6).
3. Each torus is perpendicular to each circle.

Exercise 6.6 In Chapter 4 we made a nonorientable three-manifold by gluing a room's front wall to its back wall with a side-to-side flip, while gluing the floor to the ceiling, and the left wall to the right wall, in the usual way. This nonorientable three-manifold is a product. What's it the product of? (Hint: The relevant pictures look just like Figures 6.5 and 6.6, only the gluings are different.) Is this a geometrical product? □

It's time for a brand new three-manifold with a brand new local geometry! The manifold is $S^2 \times S^1$ (read "a sphere cross a circle" or "S-two cross S-one"), but before investigating it, let's pause for a moment to see how a Flatlander might deal with $S^1 \times S^1$.

A Flatlander can't visualize $S^1 \times S^1$ directly, so he uses a mental trick. He first imagines $S^1 \times I$ as in Figure 6.7, and then he imagines the inner (circular) edge to be glued to the outer (circular) edge. If he

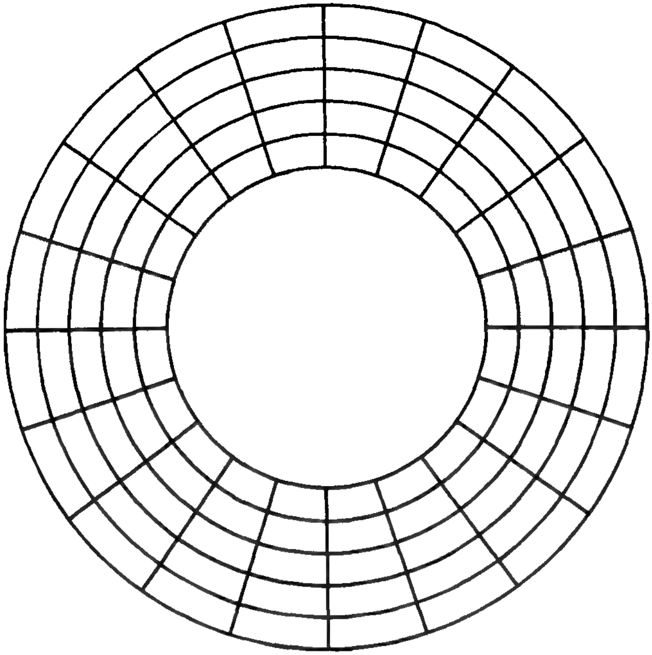


Figure 6.7 $S^1 \times I$ as imagined by a Flatlander.

wants a geometrical product, he must in addition pretend that all the circles are the same size! At first it's hard for him to pretend that these circles are the same size when in his mental image they clearly are not; but doing so does help him understand $S^1 \times S^1$, so he puts up with the contradiction until he eventually gets an intuitive feeling for what $S^1 \times S^1$ is *really* like.

Actually, the hardest thing for a Flatlander to accept is that a thread pulled taut between the points shown in Figure 6.8 would follow the apparently curved arc of one of the circles. In the more refined drawing of a Spacelander (shown on the right in Fig-

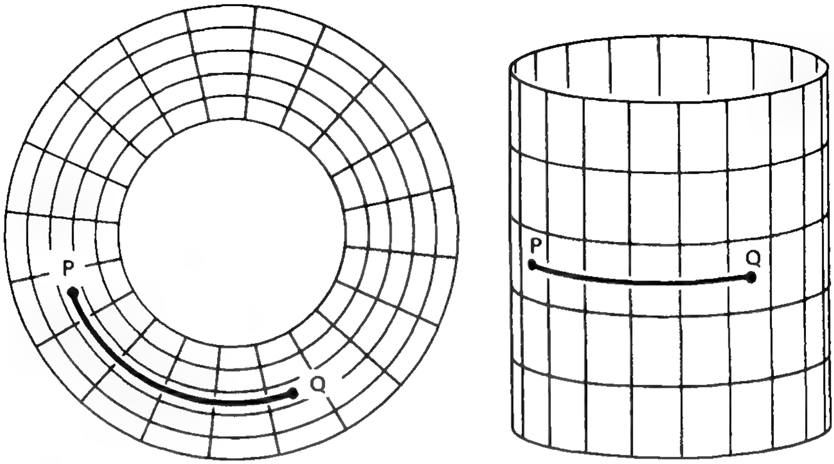


Figure 6.8 A thread pulled taut between P and Q follows the path shown.

ure 6.8) it's obvious that the thread *would* follow the circle, but to the Flatlander it seems that the thread would want to "cut across the middle" somehow.

We Spacelanders picture $S^2 \times S^1$ the same way Flatlanders picture $S^1 \times S^1$. First we visualize $S^2 \times I$ as a thickened spherical shell (Figure 6.9). To convert this shell to $S^2 \times S^1$ we imagine the inner spherical boundary to be glued to the outer spherical boundary. If we want a geometrical product things get tougher: we must pretend that the various spherical layers in the spherical shell are all the same size! (A four-dimensional being would have no trouble drawing a picture in which the layers really are the same size, but for us Spacelanders it isn't so easy.)

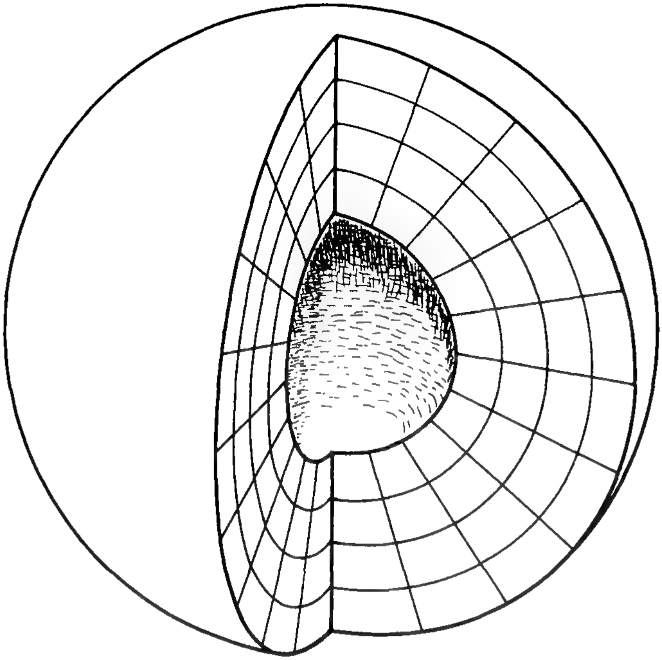


Figure 6.9 Visualize $S^2 \times I$ as a ball with a hollow center.

Figure 6.10 shows two interesting surfaces contained in a geometrical $S^2 \times S^1$. The grey one is a spherical "cross-section" of $S^2 \times S^1$. A thread pulled taut between any two points on this sphere will follow the sphere's curved surface! This would be perfectly obvious to a four-dimensional being: in her drawing there'd be no reason for the thread to bend to one side of the sphere or the other (compare Figure 6.8(b)). But to us naive Spacelanders it seems (incorrectly!) that the thread ought to "cut across the middle" somehow.

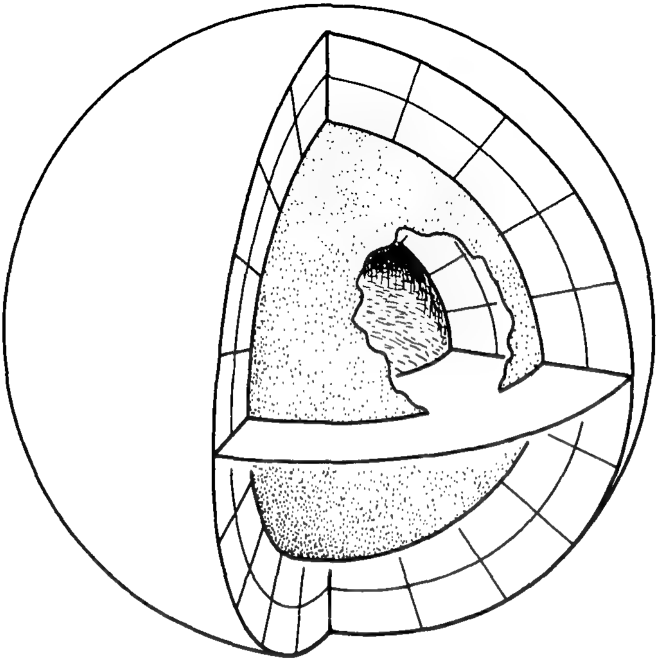


Figure 6.10 A sphere cross a circle is what you get by imagining the inner surface of $S^2 \times I$ to be glued to the outer surface. The gluing converts the horizontal ring-like surface into a torus. The grey sphere, of course, remains unaffected by the gluing.

If Gauss lived in $S^2 \times S^1$ and surveyed a triangle formed by three points on this sphere, he'd find the sum of the angles of the triangle to be greater than 180°—just like the sum of the angles of any spherical triangle (recall Figure 3.7).

The white surface shown in Figure 6.10 is a flat torus. It looks just like a Flatlander's conception of a

flat torus as based on Figure 6.7. If Gauss were to measure a triangle in this surface, he'd find its angles to add up to 180°.

We've seen that in $S^2 \times S^1$, different triangles may have different angle sums depending on how they're situated. Thus, $S^2 \times S^1$ is an example of a three-manifold which is homogeneous but not isotropic. A *homogeneous* manifold is one whose local geometry is everywhere the same. An *isotropic* manifold is one in which the geometry is the same *in all directions*. $S^2 \times S^1$ is not isotropic because some two-dimensional slices have the local geometry of a sphere, while other slices have the local geometry of a plane, as we have discovered in the preceding two paragraphs. (The distinction between homogeneous and isotropic manifolds is not readily apparent to most Spacelanders because of the peculiar fact that there are no surfaces that are homogeneous but not isotropic.)

Exercise 6.7 Find a nonorientable three-manifold that is a product and has the same local geometry as $S^2 \times S^1$. Hint: What might such a manifold be the product of? □

We now have an infinite number of three-manifolds at our disposal! Specifically, we can take each surface on the list in Exercise 5.10 and construct the product of that surface and a circle. For example, we could construct a five-holed doughnut surface cross a circle. Topologically we picture this manifold as a

thickened five-holed doughnut surface with its inside glued to its outside, just as we pictured a sphere cross a circle ($S^2 \times S^1$) as a thickened sphere with its inside glued to its outside. We won't be able to understand the geometries of these product three-manifolds until after we've investigated the geometries of surfaces (Chapters 10, 11, and 18).

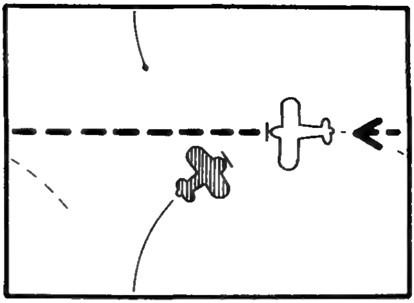
This Page Intentionally Left Blank

7

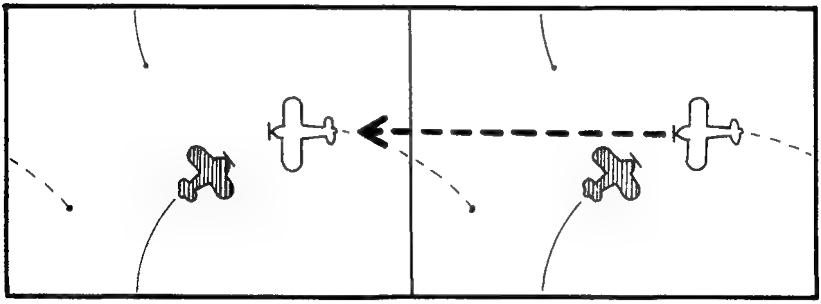
Flat Manifolds

What would you see if you lived in a closed three-manifold? This chapter answers that question for three-manifolds, such as the three-torus, that have the local geometry of "ordinary" three-dimensional space (i.e. flat three-manifolds).

First let's consider what a two-dimensional biplane pilot sees as she flies about in the flat torus sky of Chapter 3. If she looks towards an "edge of the screen," she'll find she's looking at herself from behind! (See Figure 7.1.) She can also see herself by looking along a diagonal, as illustrated in Figure 7.2.



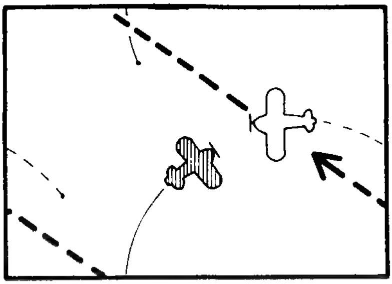
The pilot can
see herself.



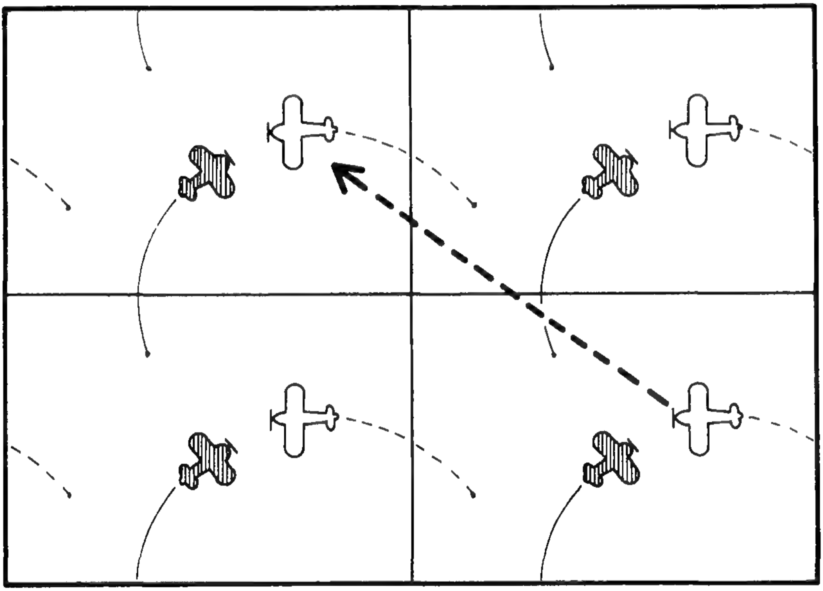
How the situation
appears to her.

Figure 7.1 A biplane pilot in a flat torus.

In fact she can see herself in infinitely many directions! The flat torus sky looks to her like a plane with infinitely many copies of her in it! (Figure 7.3.) Of course if an enemy biplane is chasing her she'll see infinitely many copies of it as well. These pictures are analogous to the pictures we used in Chapter 4 to analyze Klein bottle tic-tac-toe games (Figure 4.6).



She sees herself by looking along a diagonal.



How it appears to her.

Figure 7.2 The biplane pilot sees herself in a different direction.

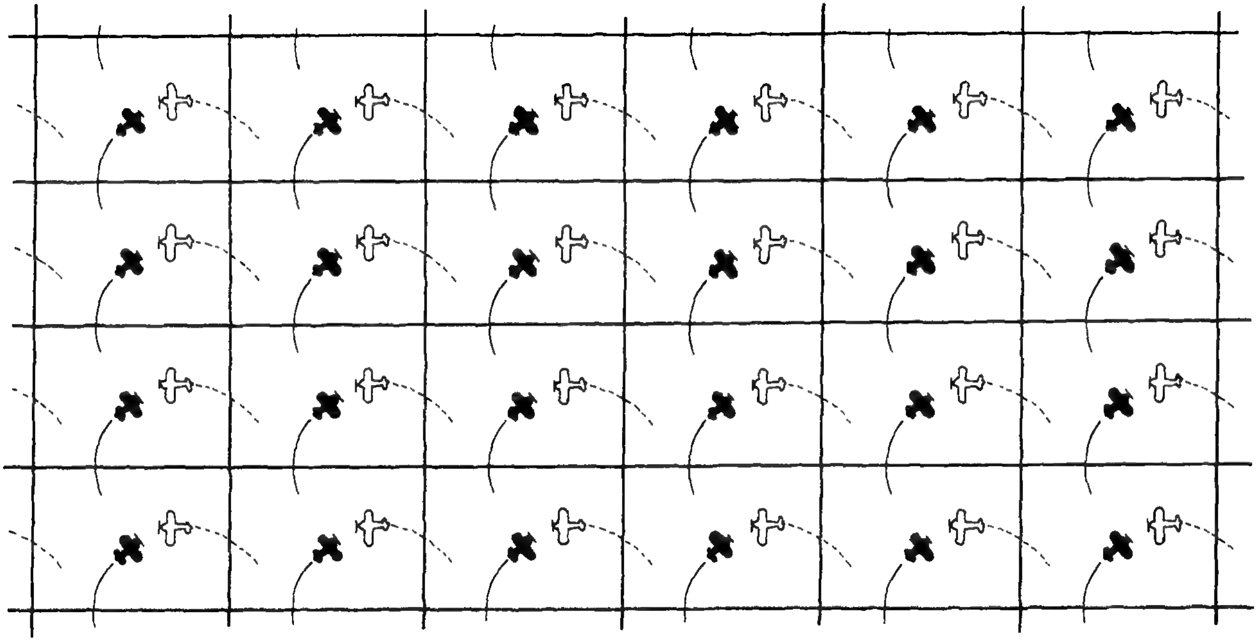


Figure 7.3 The biplane pilot's view of her world.

Exercise 7.1 Draw a picture (analogous to Figure 7.3) showing how the sky would appear to a biplane pilot flying in a Klein bottle. □

Exercise 7.2 Analyze the *torus* tic-tac-toe game of Exercise 2.2 using the method of Figure 4.6. □

Now let's make some analogous drawings of three-manifolds. Figure 7.4 shows the view inside the three-torus described in Chapter 2. This three-torus consists of a living room with its left wall glued to its right wall, its front wall glued to its back wall, and its floor glued to its ceiling.

Figure 7.5 shows the view inside a different three-torus: this one is made from a cube containing a smaller cube with colored faces (Figure 7.6).

DO-IT-YOURSELF COLORING

Get ahold of some colored pens or pencils and use them to color the faces of the little cubes in Figures 7.5 and 7.6 according to the code R = Red, O = Orange, Y = Yellow, G = Green, B = Blue, V = Violet. You'll eventually need to color Figures 7.7, 7.10, 7.11, and 7.12, and the figures for Exercise 7.3, according to the same code.

It's fun to imagine the view inside other three-manifolds made from the (big) cube of Figure 7.6. For

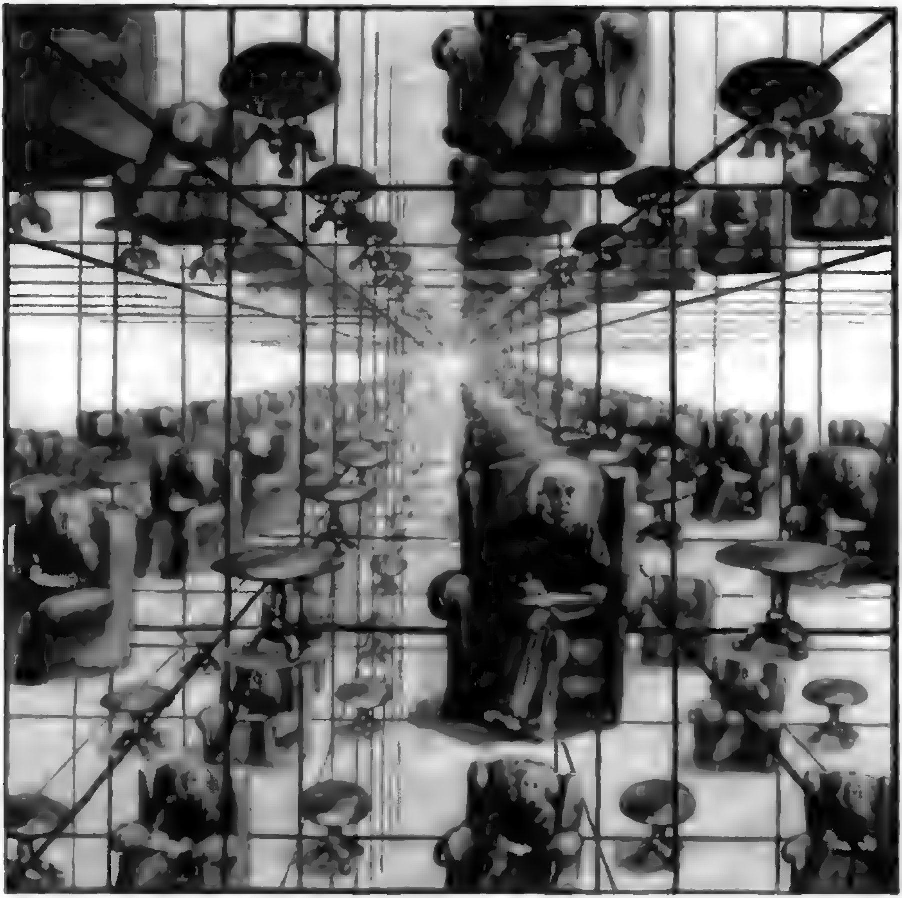


Figure 7.4 When opposite walls of a room are "glued" to make a three-torus, the view looks like this.

example, consider making $K^2 \times S^1$ from this cube. ($K^2 \times S^1$ is the nonorientable three-manifold described in Chapter 4—it's made by gluing a cube's top and bottom, and left and right sides, in the usual way, but gluing the front to the back with a side-to-side flip.)

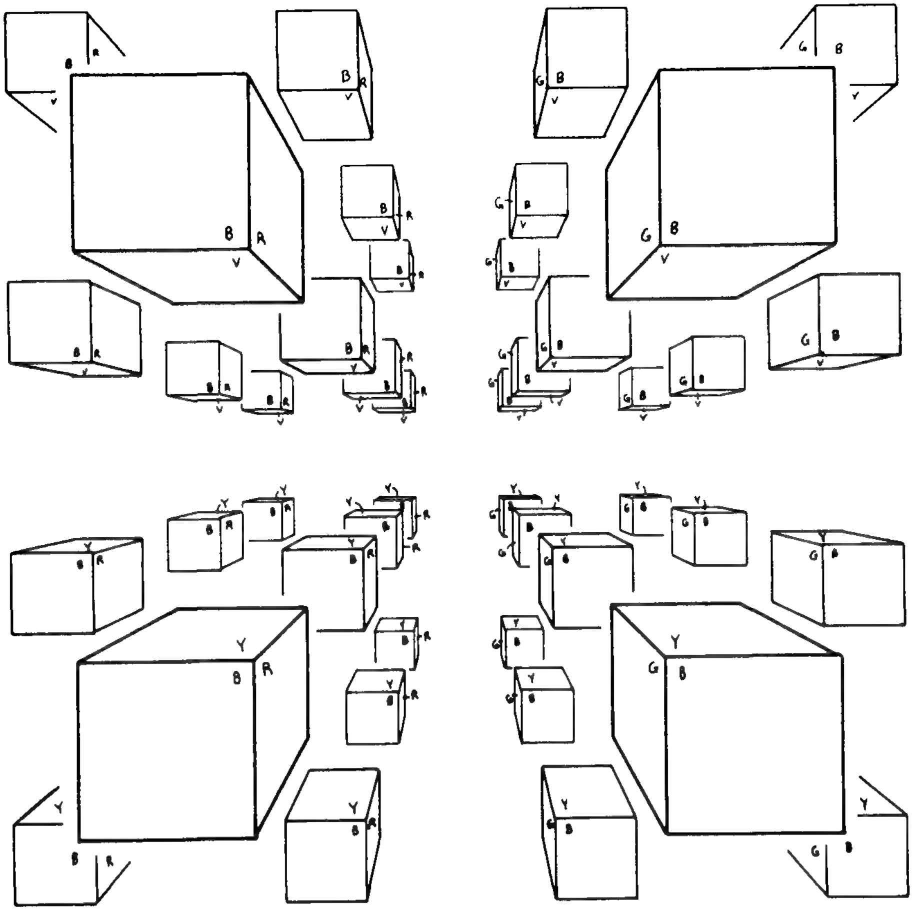


Figure 7.5 The view inside a three-torus containing a single small cube.

Color Figure 7.7 as indicated by the code to see the view inside $K^2 \times S^1$. *IMPORTANT NOTE: On the little cube opposite faces have complementary colors: red is opposite green, blue is opposite orange, and yellow is opposite violet.* Imagine yourself to be sitting on one

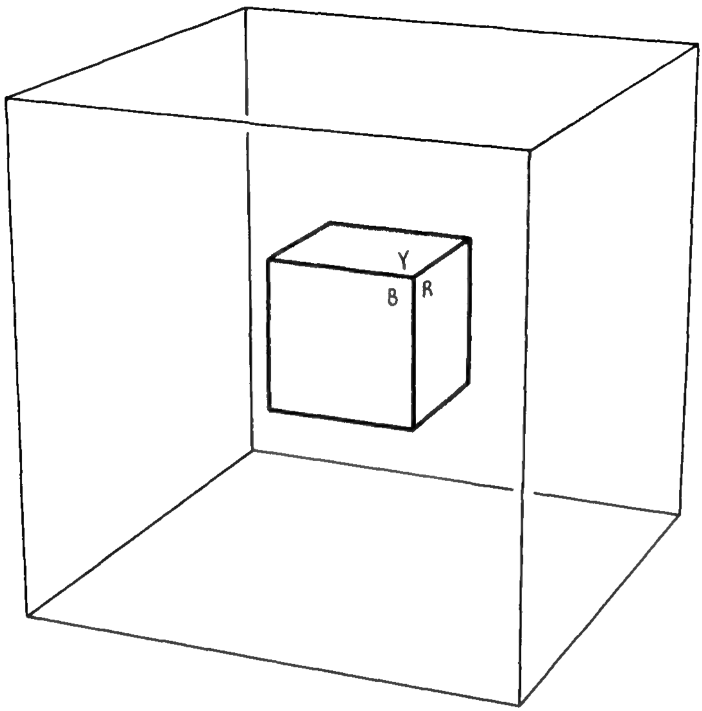


Figure 7.6 If you glue opposite faces of this (big) cube, the view inside the resulting three-torus will be as shown in Figure 7.5.

of the cubes in Figure 7.7. When you look up, down, to the left or to the right you see other colored cubes positioned just as the cubes in Figure 7.5 are. But the cubes immediately in front of you and immediately behind you appear to have undergone a side-to-side mirror reversal (thus interchanging the red and green faces); this is because the front and back faces of the big cube were glued with a side-to-side flip. The view in $K^2 \times S^1$ is analogous to the drawing of a Klein bottle

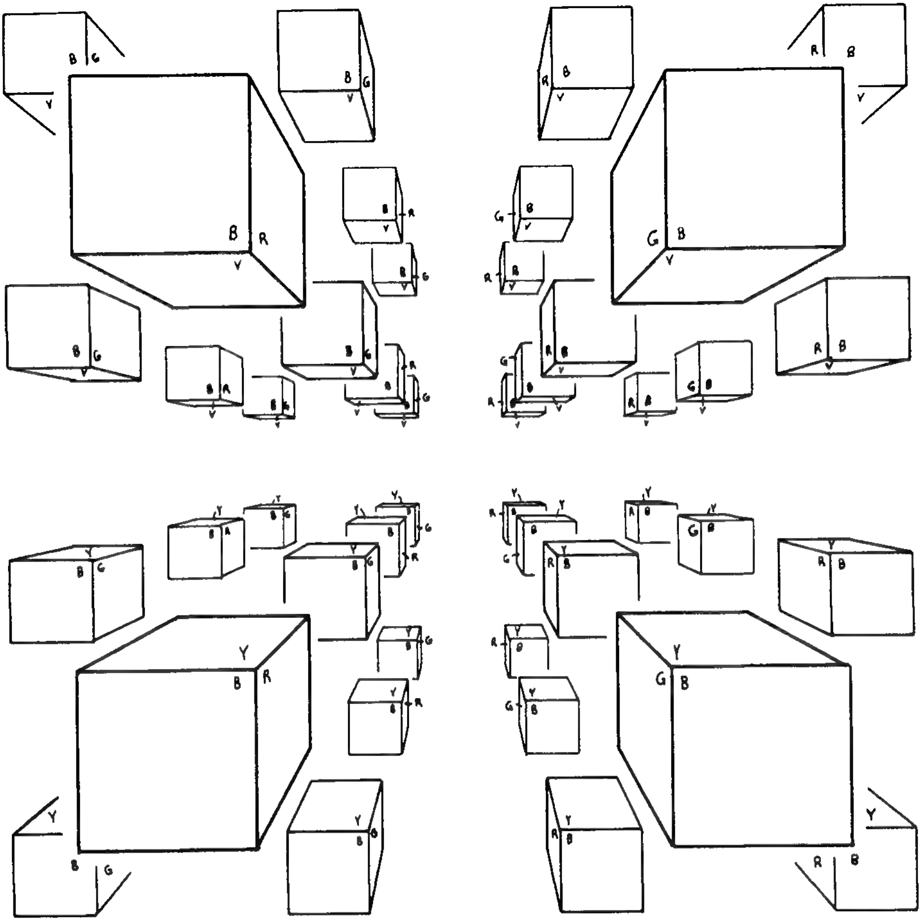


Figure 7.7 The view inside $K^2 \times S^1$.

sky you made in Exercise 7.1. You might want to pause for a moment to think this through.

Exercise 7.3 Here's a new three-manifold, the "quarter turn manifold." Start with a cube like the one in Figure 7.6, glue its front, back, left and right sides as

if you were making a three-torus, but then glue its top to its bottom with a quarter turn (Figure 7.8 shows how). Make a copy—by hand or by photocopy machine—of Figure 7.9 and color it to show the view inside this manifold.

You might also want to illustrate the views in a "half turn manifold" and a "three quarters turn manifold." Compare the view in the three quarters turn manifold to the view in the original quarter turn manifold; can you intrinsically tell one from the other? □

In the preceding exercise you colored the view inside the quarter turn manifold based on a knowledge of how the cube's faces were glued. You can also reverse this process: given a colored illustration of the view inside a manifold, you can deduce what the gluings are that produce it. For example, in Figure 7.7,

1. Each little cube is the mirror image—via a left-to-right reversal—of the cube immediately behind it. This tells you that the big cube's front and back faces are glued with a side-to-side flip.
2. Each little cube is identical to the cube above it. This tells you that the big cube's top and bottom faces are glued normally.
3. Similarly, each little cube is identical to the cube to its left. This tells you that the left and right faces are glued normally.

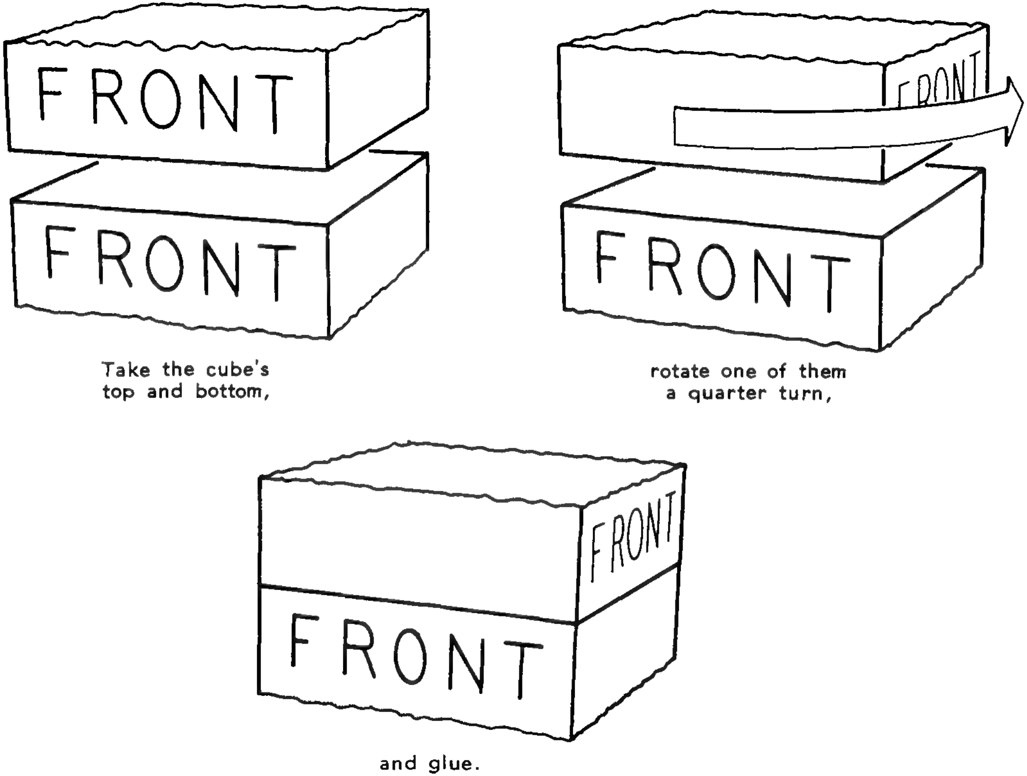


Figure 7.8 How to glue the top of a cube to the bottom with a quarter turn (local intrinsic picture).

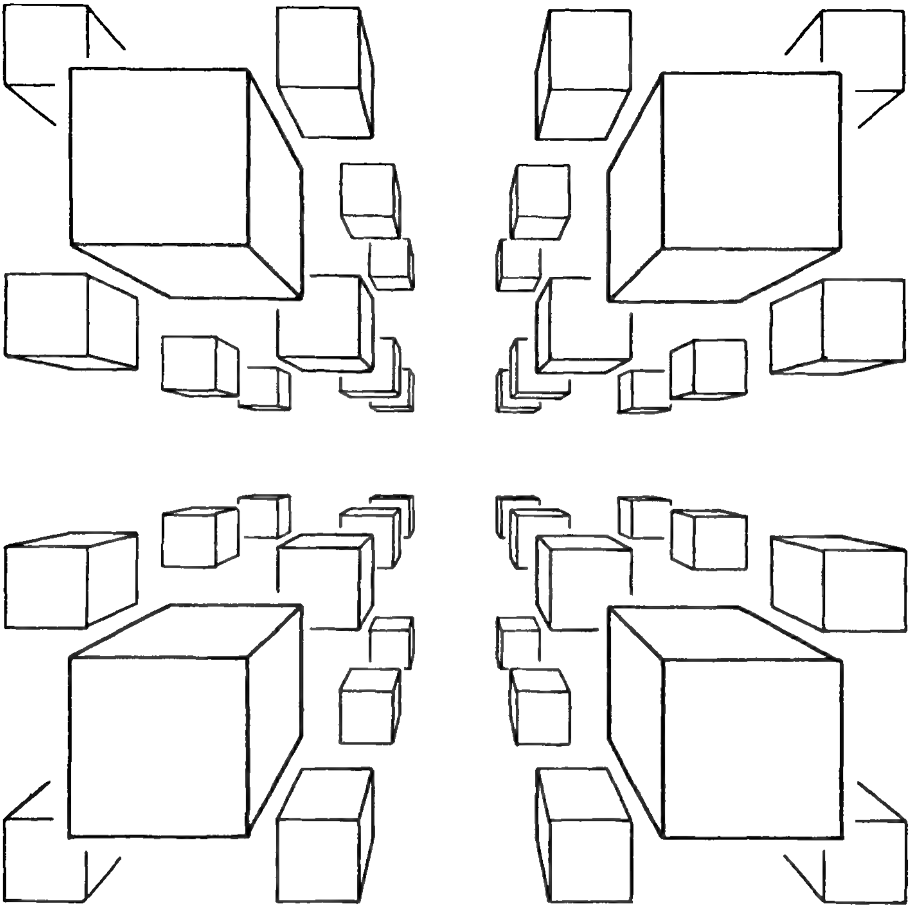


Figure 7.9 Make several copies of this figure and color them to illustrate the view inside each manifold of Exercise 7.3.

If you apply similar reasoning to your colored illustration of the view inside the quarter turn manifold, you can "rediscover" that the cube's top and bottom are glued with a quarter turn.

Exercise 7.4 Color Figure 7.10 as indicated by the code. How should the (big) cube of Figure 7.6 be glued to produce this view? For more practice, color Figures 7.11 and 7.12, and decide what gluings will produce these views. □

Exercise 7.5 Which of the manifolds from Exercises 7.3 and 7.4 are orientable and which are nonorientable? (Recall that to determine orientability you have to check whether crossing a face of the cube can bring you back from the opposite face mirror reversed.) □

Two comments are in order concerning these three-manifolds made from cubes. First, each has the local geometry of ordinary three-dimensional Euclidean space. In other words, each is flat. Second, if you suddenly found yourself in one of these manifolds—but you didn't know which one it was—you could easily check for orientability by raising your right hand. If all the other copies of yourself raised their right hands you'd know the manifold was orientable; if any of them raised their left hands you'd know it was nonorientable. With more detailed observations you could determine the nature of the manifold exactly (as in Exercise 7.4, only now you yourself take the place of the colored cube).

Not all three-manifolds are made from cubes, and not all surfaces are made from squares. For example, consider a regular hexagon whose opposite edges are

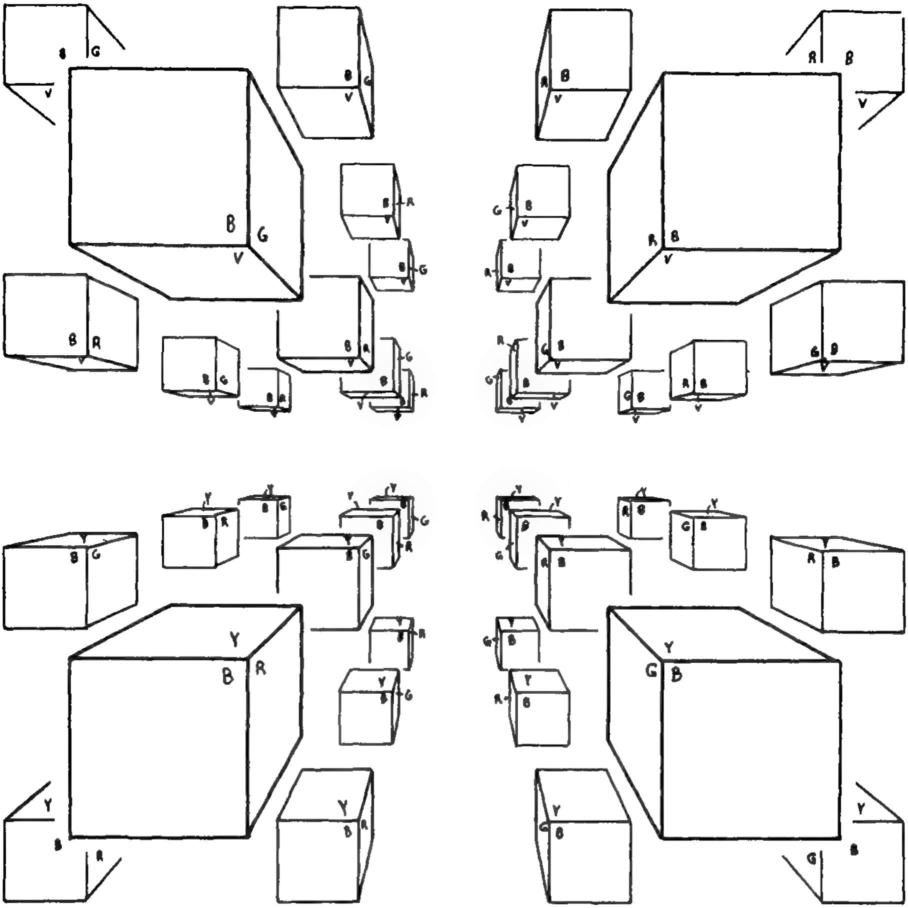


Figure 7.10 What gluing of the big cube produces this view?

abstractly glued. If you are willing to deform the hexagon you can physically carry out the gluings in three-dimensional space (Figure 7.13). You will find that the surface has the same global topology as a torus. Because the surface is made from a hexagon and has the topology of a torus, it is called a *hexagonal torus*. Its

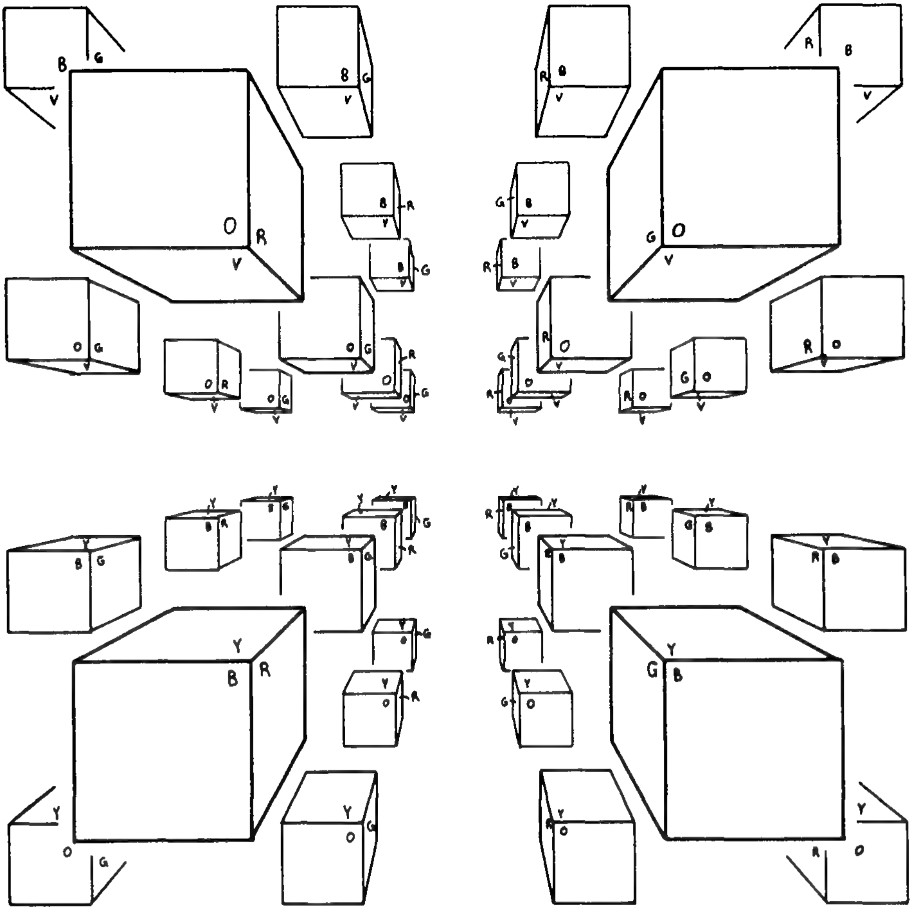


Figure 7.11 What gluing of the big cube produces this view?

local geometry is flat because the hexagon is flat (geometrically you should think of the hexagonal torus as a hexagon with abstractly glued edges, rather than as a doughnut surface).

Exercise 7.6 Imagine a biplane pilot flying about in a hexagonal torus. Draw a picture analogous to Figure

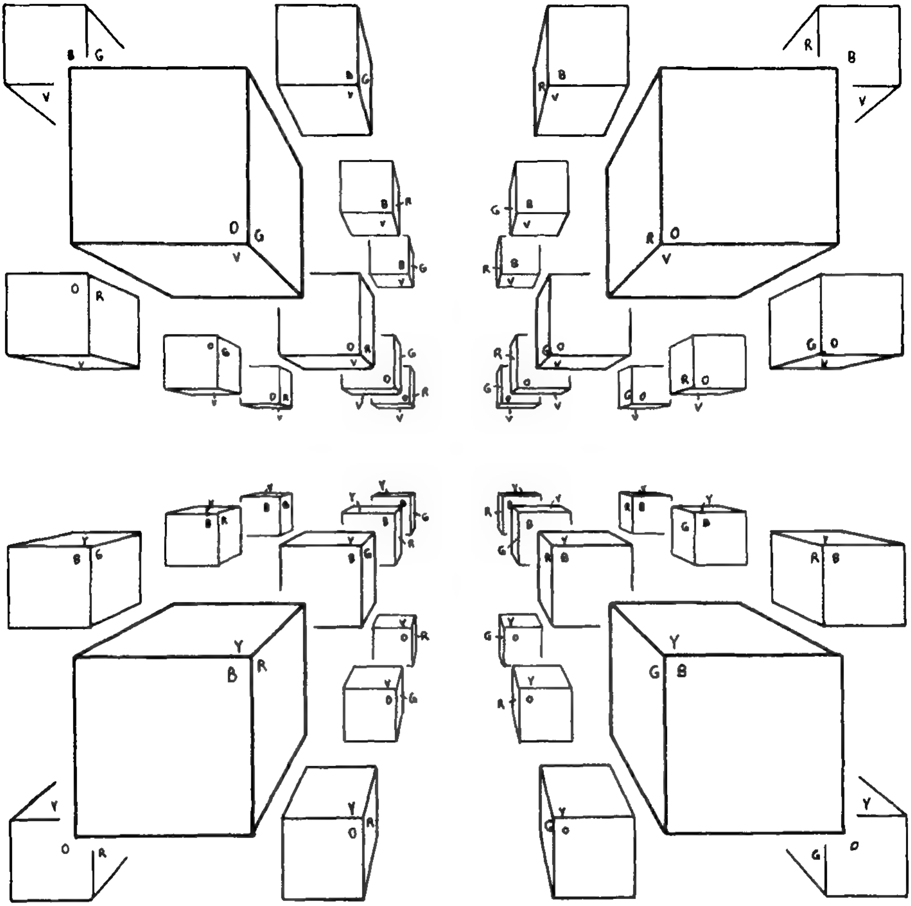


Figure 7.12 What gluing of the big cube produces this view?

7.3 showing how the hexagonal torus appears to the pilot. What if a second biplane is flying around with her? □

For the most part the hexagonal torus is very similar to the usual flat torus. One way that it's different,

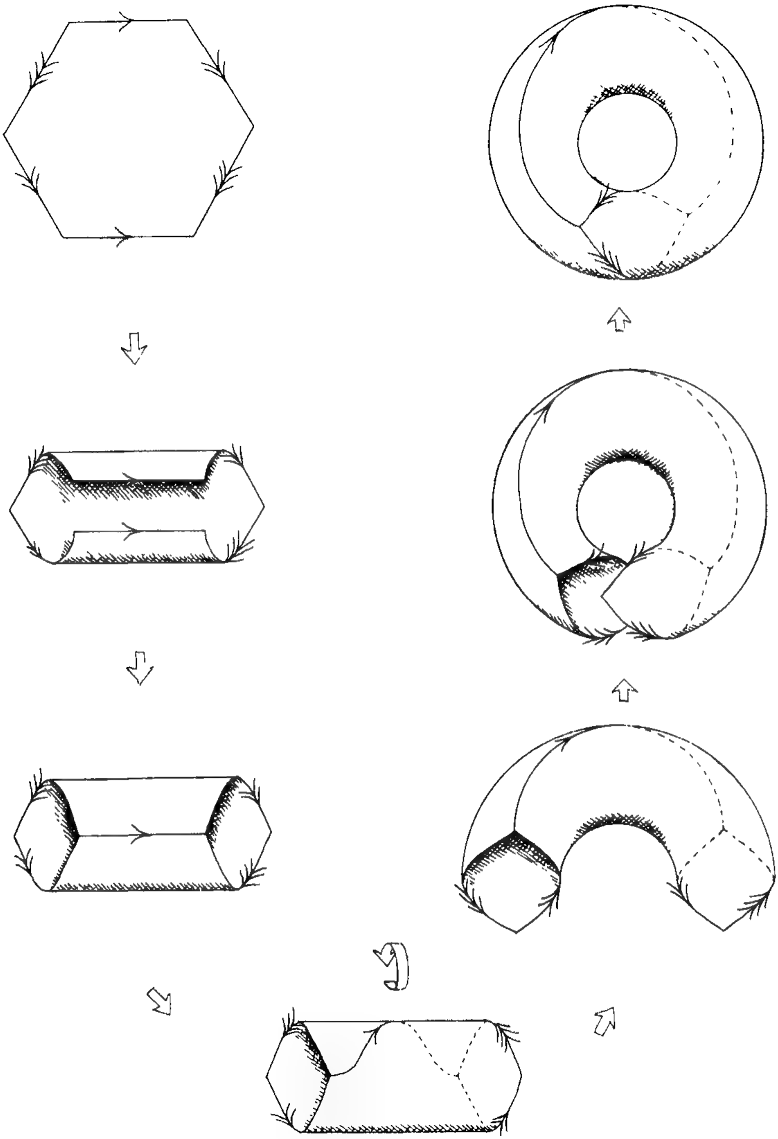


Figure 7.13 To physically glue together opposite edges of a hexagon, you must deform the hexagon into the shape of a doughnut surface. A hexagon with abstractly glued edges therefore has the same global topology as a torus.

however, is that all six corners of the hexagon do not get glued to a single point. Compare Figure 7.13 to Figure 3.3. In Figure 3.3 all four corners of the square meet at a single point in the surface, whereas in Figure 7.13 the hexagon's six corners meet in two groups of three corners each.

Later on it will be important to know how a polygon's corners fit together, and in many cases we won't be able to physically carry out the gluings in three-dimensional space. Fortunately we can always tell how the corners fit simply by studying the polygon itself. For example, imagine a Flatlander in the ordinary flat torus who decides to go for a walk around the point at which the square's four corners meet. (Intrinsically he could never locate such a point, but for the sake of argument pretend he is taking a walk around just that point.) Ignore the lower picture in Figure 7.14, and follow his progress in the upper picture. He begins in the lower right hand corner of the square. He passes through the right edge of the square into the lower left hand corner. From there he moves through the bottom edge into the upper left hand corner, then through the left edge into the upper right hand corner, and finally through the top edge to get home. By following his progress in the square we have deduced that all four corners meet at a single point. In contrast, a Flatlander going for a walk in a hexagonal torus would visit only three corners before returning home, either corners 1, 2, and 3 or corners a, b, and c, depending on where he started (see Figure

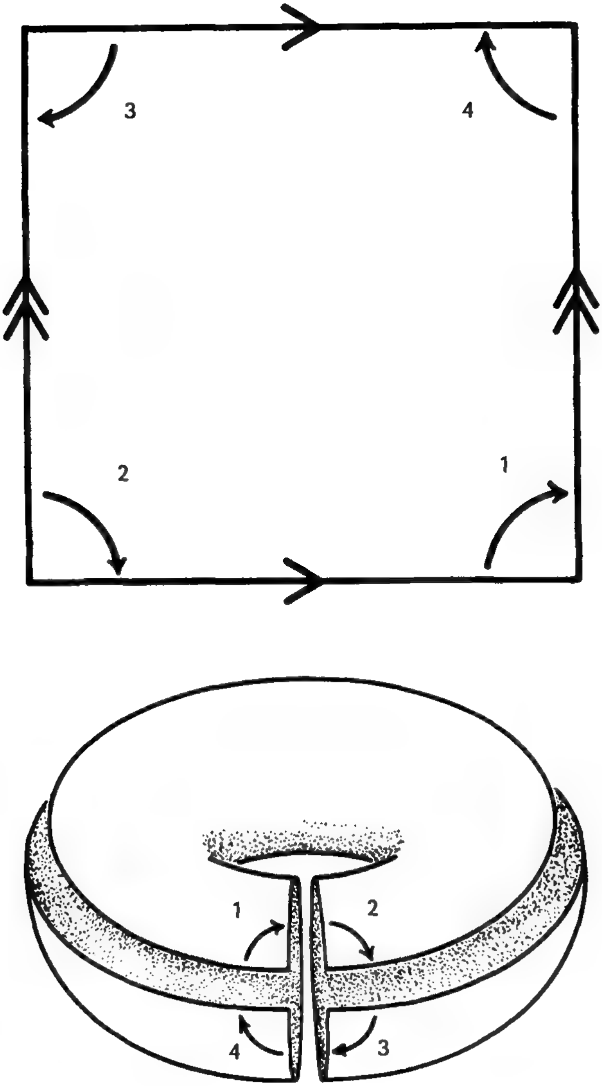


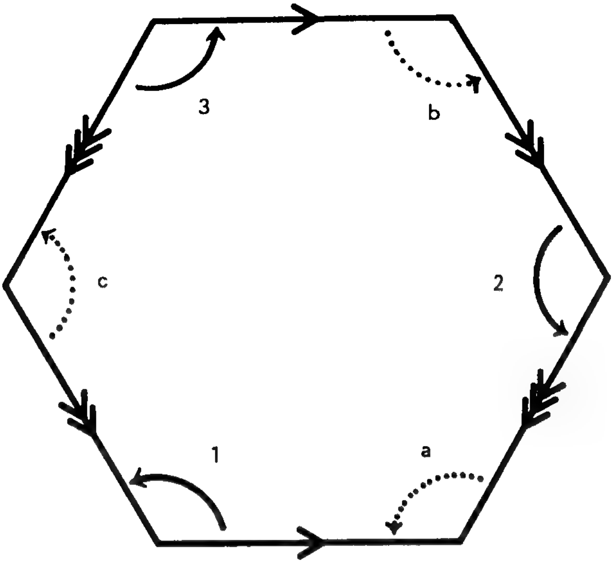
Figure 7.14 A Flatlander walking in a small circle can visit all four corners.

7.15). Now let's use this "walking around corners" technique to study a new surface. Consider a square whose edges are glued as indicated by the arrows in Figure 7.16 (note: *both* pairs of edges are glued with a flip). A Flatlander starting in the lower right hand corner passes through the right edge into the *upper* left hand corner; from there he passes through the top edge to get home. Similarly, a Flatlander starting in the lower left hand corner goes to the upper right hand corner and then home. This tells us that the square's corners meet in two groups of two corners each. But something funny is going on here: two corners won't fit together properly! Figure 7.17 shows that when two 90° corners come together you get a "cone point." In general you will get a cone point whenever the sum of the angles at a point is less than 360°.

Exercise 7.7 Figure 7.18 shows three surfaces, each a hexagon with edges glued as indicated by the arrows. Use the "walking around corners" technique described above to decide how each hexagon's corners fit together. Which of these surfaces have cone points? □

Exercise 7.8 Which of the surfaces in the preceding exercise are orientable? □

Exercise 7.9 Draw a picture analogous to Figure 7.3 showing a biplane pilot's impression of the second surface in Figure 7.18. □



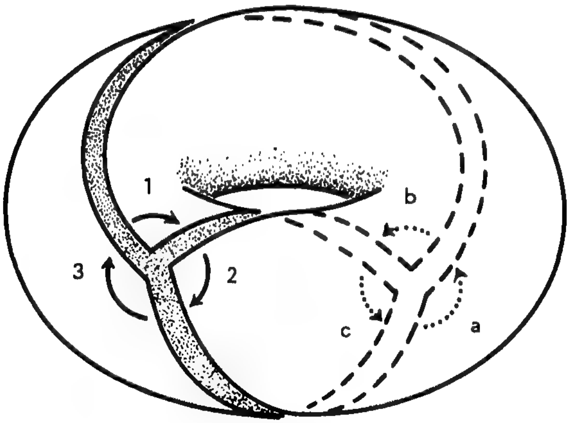


Figure 7.15 A Flatlander walking in a small circle visits a group of three corners.

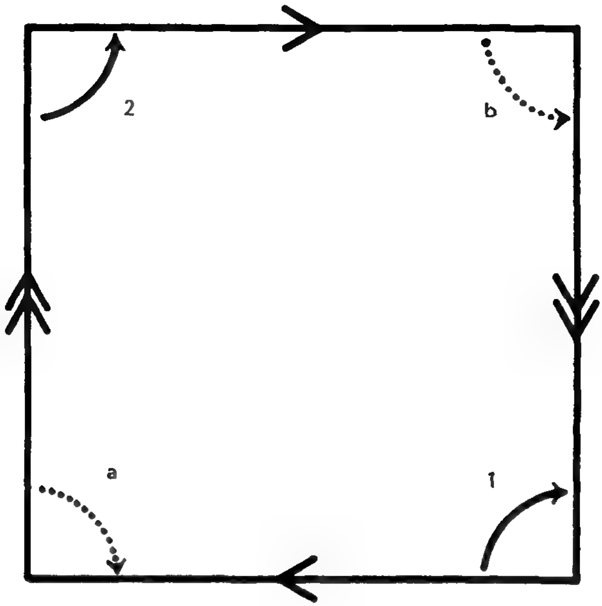


Figure 7.16 Glue each pair of opposite sides with a flip. A Flatlander walking in a small circle can visit only two corners.

Exercise 7.10 Do a hexagonal torus and an ordinary flat torus have the same local topology? The same local geometry? The same global topology? The same global geometry? □

At this point I'd like to clear up a definitional matter. When we say a surface has a flat geometry we mean that its local geometry is flat (i.e. Euclidean) at *all* points. Cone points are not allowed. For example, the surface defined in Figure 7.16 is not considered flat, even though it has a flat local geometry at all points except two. In contrast, a hexagonal torus is

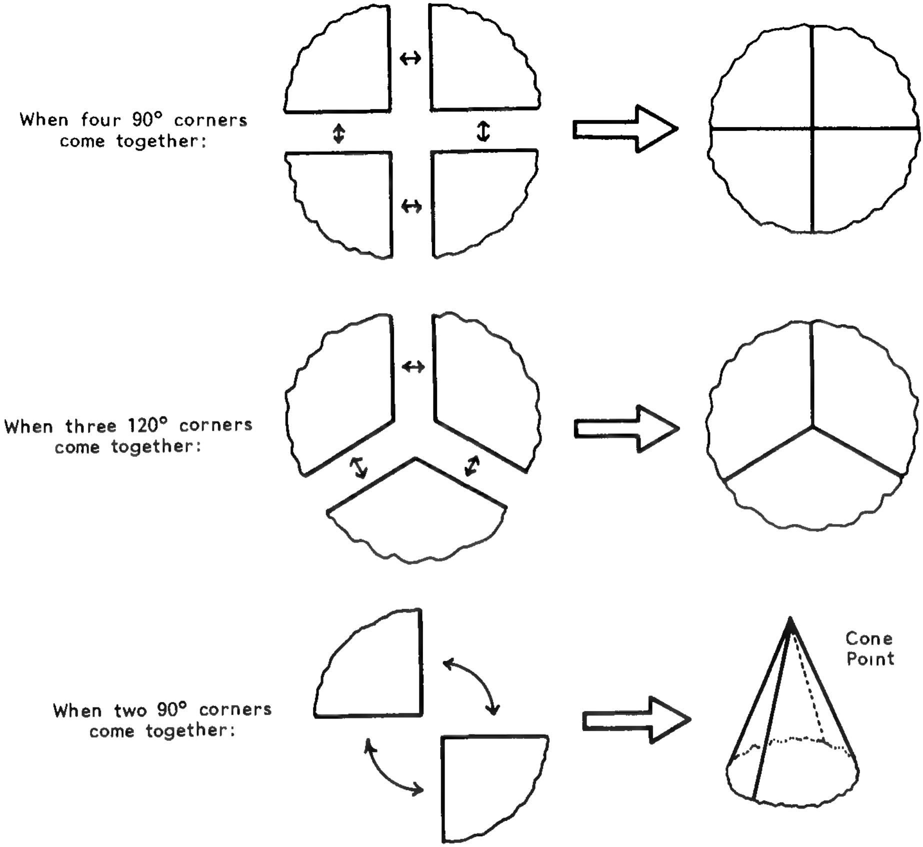


Figure 7.17 If the corners add up to less than 360° you'll get a "cone point."

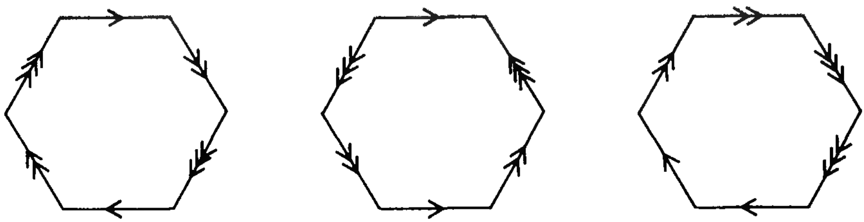
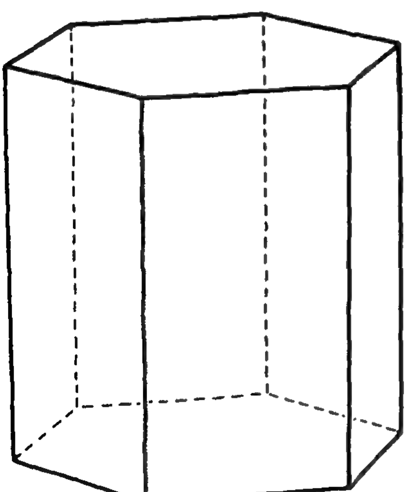


Figure 7.18 Which of these surfaces have cone points?

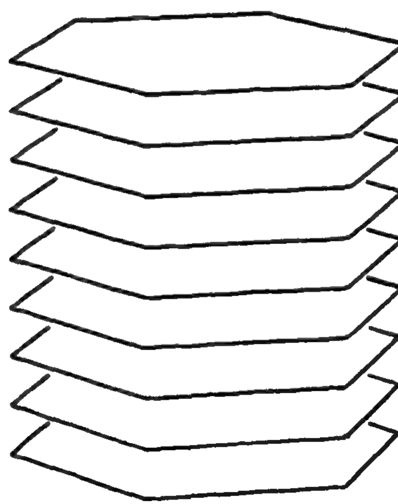
considered flat because it has a flat local geometry at all points, without exception.

Exercise 7.11 Which of the surfaces of Exercise 7.7 are flat? □

The remaining three-manifolds in this chapter will be made from a hexagonal prism like the one shown in Figure 7.19(a). When you glue the opposite side faces of this prism each horizontal layer (see Figure 7.19(b)) becomes a hexagonal torus and you get a hexagonal torus cross an interval. If you then glue the bottom of the prism to the top you'll wind up with a hexagonal torus cross a circle, which is new form of



(a) The basic prism



(b) The prism thought of as a hexagon cross an interval

Figure 7.19 A hexagonal prism is useful for making three-manifolds.

three-dimensional torus. It's called a *hexagonal three-torus*.

Exercise 7.12 Do the hexagonal three-torus and the usual three-torus have the same local geometry? The same global topology? The same global geometry?

Imagine yourself inside a hexagonal three-torus. You see other copies of yourself. How are they arranged? □

Exercise 7.13 Use a hexagonal prism to construct a "1/3 turn manifold" and a "1/6 turn manifold" analogous to the quarter turn manifold of Exercise 7.3. If you lived in one of these manifolds, how would the other copies of yourself be arranged? □

At first glance it's not at all obvious that the hexagonal three-torus and the usual three-torus are topologically the same (i.e. have the same global topology). Similarly, in Chapter 5 it was a challenge to understand that $P^2 \# P^2 \# P^2$ and $T^2 \# P^2$ were the same. This leaves one in doubt as to whether other seemingly different manifolds might also be topologically the same. Around the turn of the century mathematicians invented various tools to deal with this problem. For example, the Euler number (Chapter 12) can conclusively decide whether or not two surfaces are the same topologically. More complicated tools are required for three-manifolds, such as the "homology groups" and the "fundamental group" (see Chapters 5 and 8 of Armstrong's *Basic Topology*, or any algebraic

topology book). Fortunately in practice manifolds that look different usually are.

You can explore the 3-manifolds of this chapter using interactive 3-D graphics software available for free at *www.northnet.org/weeks/SoS*.

8

Orientability vs. Two-Sidedness

A Möbius strip is nonorientable. In addition, the Möbius strip in Figure 8.1 has only one side, as Escher's ants are demonstrating. In general a surface lying in a three-manifold is called *one-sided* if a (three-dimensional!) ant can go for a walk along it and come back on the opposite side of the surface from which she started. It's often thought that nonorientability and one-sidedness are just two aspects of the same prop-

---

This chapter is dedicated to Bob Messer, who showed me a two-sided Möbius strip at a time when I hadn't the slightest idea that such a thing could possibly exist.

Figure 8.1 Escher's ants demonstrate that this Möbius strip has only one side. (Möbius Strip II by M. C. Escher, copyright © Beeldrecht, Amsterdam/VAGA, New York. Collection Haags Gemeentemuseum—The Hague, 1985.)

erty, but this is not the case, as you can see from the examples in Exercise 8.1.

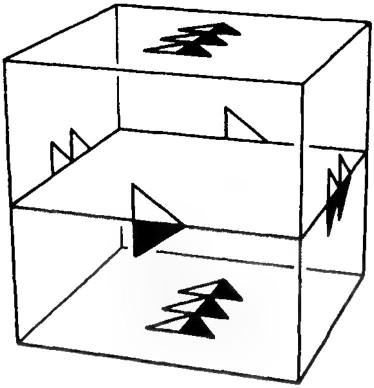
A surface (lying in a three-manifold) that is not one-sided is called *two-sided*, because in this case the surface has two separate sides, and a (three-dimensional) ant walking on the surface can never get from one side to the other.

Exercise 8.1 Figure 8.2 shows four ways to construct a three-manifold by gluing opposite faces of a cube. In each case the cube contains a square which gets glued to form a surface in the three-manifold. For each example identify the surface and say whether it's orientable or nonorientable, and whether it's one-sided or two-sided. To determine orientability you have to check whether a (two-dimensional) Flatlander can go for a walk *in* the surface and return to his starting point mirror reversed. To determine "sidedness" you have to check whether a (three-dimensional) ant can go for a walk *on* the surface and return to her starting point on the opposite side of the surface.

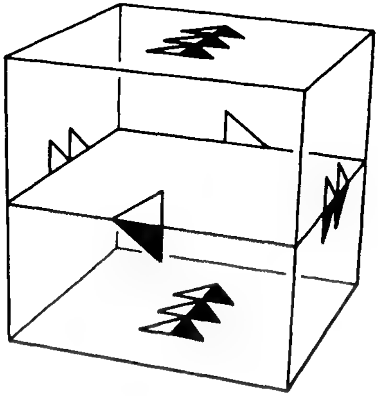
Use the four examples to fill in the following table:

	orientable	nonorientable
one-sided		
two-sided		

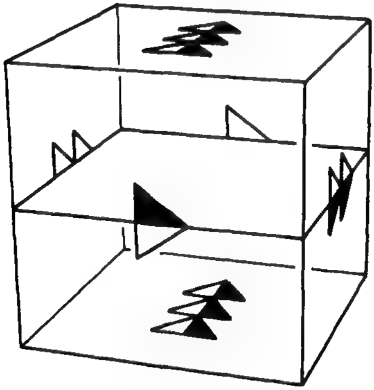
□



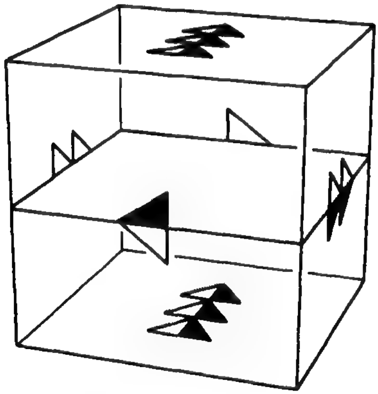
(1) Glue opposite faces in the usual way to form a 3-torus.



(2) Glue the top to the bottom, and the left side to the right side, in the usual way, but glue the front to the back with a side-to-side flip. This 3-manifold is $K^2 \times S^1$.



(3) Glue the top to the bottom, and the left side to the right side, in the usual way, but glue the front to the back with a top-to-bottom flip. This 3-manifold is also $K^2 \times S^1$.



(4) Glue the top to the bottom, and the left side to the right side, in the usual way, but glue the front to the back with both a side-to-side flip *and* a top-to-bottom flip. The two flips combined are the same as a 180° rotation, and the manifold is the half-turn manifold of Exercise 7.3.

Figure 8.2 Four three-manifolds, each containing a surface.

Exercise 8.2 Is sidedness an intrinsic or an extrinsic property of a surface? What about orientability? Are all Klein bottles nonorientable? Are they all one-sided? □

Exercise 8.3 Construct a three-manifold containing a two-sided Möbius strip. Just for fun, draw some red ants on one side of the Möbius strip and some black ants on the other. Note that the red ants and black ants can never meet (at least not without crawling around the edge of the Möbius strip!). □

One's initial impression that nonorientability and one-sidedness are just different aspects of a single property is not entirely groundless. In an *orientable three-manifold*, every two-sided surface is orientable and every one-sided surface is nonorientable. To see why, pretend that you are standing with a friend on a surface in an orientable three-manifold. Your friend decides to go for a walk. Because you're in an orientable three-manifold you know he can't possibly come back mirror reversed. Assume for the moment that the surface you're standing on is one-sided, and that your friend comes back from his walk on the side opposite from where you're standing. Also assume that the surface is transparent so you can look through and see the bottoms of his feet. (In this manifold, gravity is directed towards the surface, so neither of you falls off!) Because your friend is on the other side of the surface, his footprints appear mirror reversed (Figure 8.3). If you study just his footprints (ignore his three-

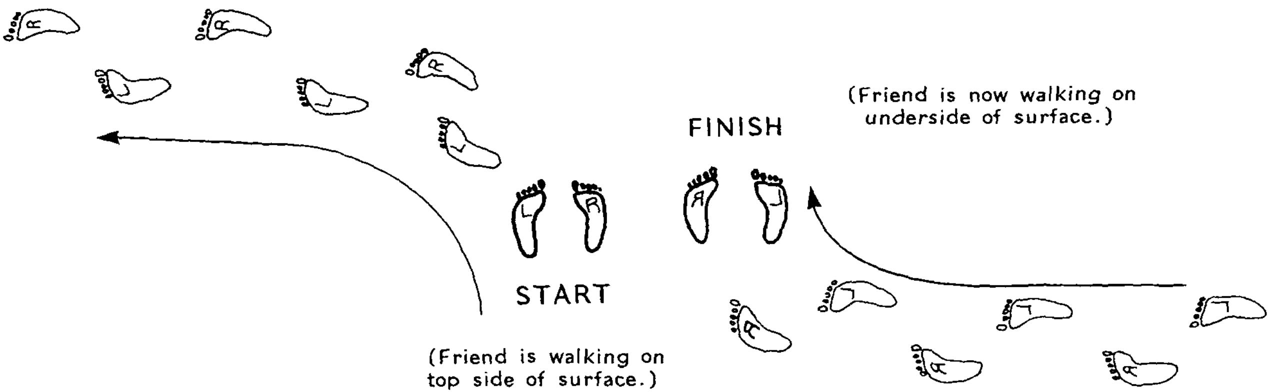


Figure 8.3 Going for a walk on a one-sided surface in an orientable three-manifold. At the end of the walk your friend *isn't* mirror-reversed; but because he's now standing on the *other* side of the surface, his footprints appear mirror-reversed to you.

dimensional body) you'll see that they trace out an orientation reversing path in the surface. Therefore this one-sided surface is nonorientable.

If, on the other hand, the surface were two-sided, then your friend would come back from his walk on the same side of the surface he started on. Because the three-manifold is orientable, he'd be his normal self rather than his mirror image. In particular, his footprints would not be mirror reversed. This shows that no matter where he walks, his footprints will never trace out an orientation reversing path in the surface. In other words, this two-sided surface is orientable.

This Page Intentionally Left Blank

Part II

Geometries on Surfaces

This Page Intentionally Left Blank

9

The Sphere

In a sequel to *Flatland*, an octagon by the name of Mr. Puncto surveys some rather large triangles and finds that each has angles that add up to slightly more than 180°! This discovery excites Mr. Puncto, and, after double checking his data, he makes the discovery public. Unfortunately neither the civil authorities nor the scientific establishment shares his excitement. They suspect he is merely inventing excuses to explain some errors in his measurements, and they dismiss him from his job. The true explanation is that the Flatlanders are living not in a plane but on a sphere,

and on a sphere the angles of a triangle really do add up to more than 180°, as we shall soon see. This incident and much more is described in the book *Sphereland* by Dionys Burger. I heartily recommend *Sphereland* to all readers of the present book (just don't be put off by the rather dull summary of *Flatland* at the beginning).

A triangle drawn on a sphere is called a *spherical triangle*. Each side of a spherical triangle is required to be a geodesic; that is, it is required to be intrinsically straight in the sense that a Flatlander on the sphere would perceive it as bending neither to the left nor to the right. A side of a spherical triangle is thus an arc of a so-called great circle (see Figure 9.1).

From now on we will *measure all angles in radians*, to facilitate easier comparison of angles and areas (in a minute you'll see how and why we want to do this). Recall that π radians = 180°, $\pi/2$ radians = 90°, etc. Except when specified otherwise, we will henceforth assume that all spheres are unit spheres, i.e. they all have radius one.

Exercise 9.1 For each spherical triangle in Figure 9.2 compute (1) the sum of its angles in radians, and (2) its area. To compute the areas, use the fact that the unit sphere has area 4π. For example, the first triangle shown occupies $\frac{1}{8}$ of the sphere, so its area is $(4\pi)/8 = \pi/2$.

Find a formula relating a spherical triangle's angle-sum to its area. This formula appeared in 1629 in

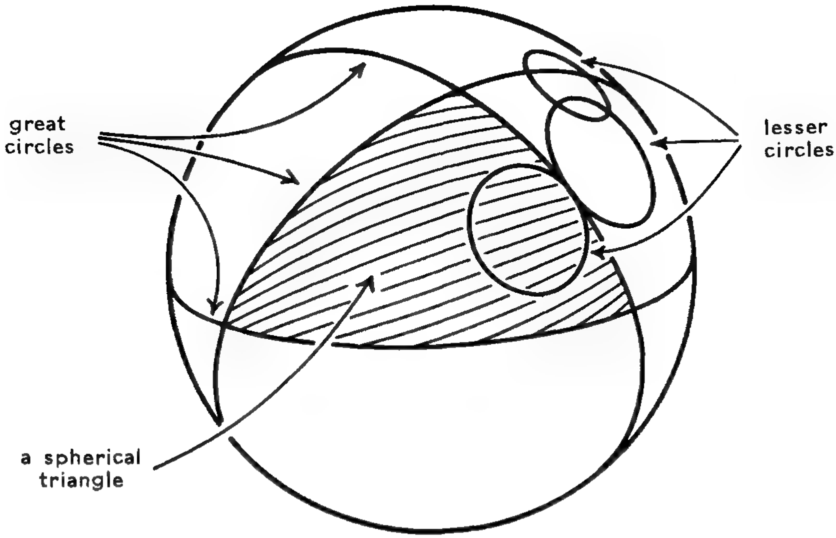


Figure 9.1 On any sphere, the *great circles* are those circles that are as big as possible. A great circle appears straight to a Flatlander on the sphere. By contrast, any lesser circle appears to bend to one side or the other.

the section "De la mesure de la superfice des triangles et polygones sphericques, nouvellement inventee Par Albert Girard" of the book *Invention nouvelle en L'Algebre* by Albert Girard.

You should try to find the formula before reading on, because the following paragraphs give it away. □

Exercise 9.2 What is the area of a spherical triangle whose angles in radians are $\pi/2$, $\pi/3$, and $\pi/4$? What is the area of a spherical triangle with angles of 61°, 62°, and 63°? □

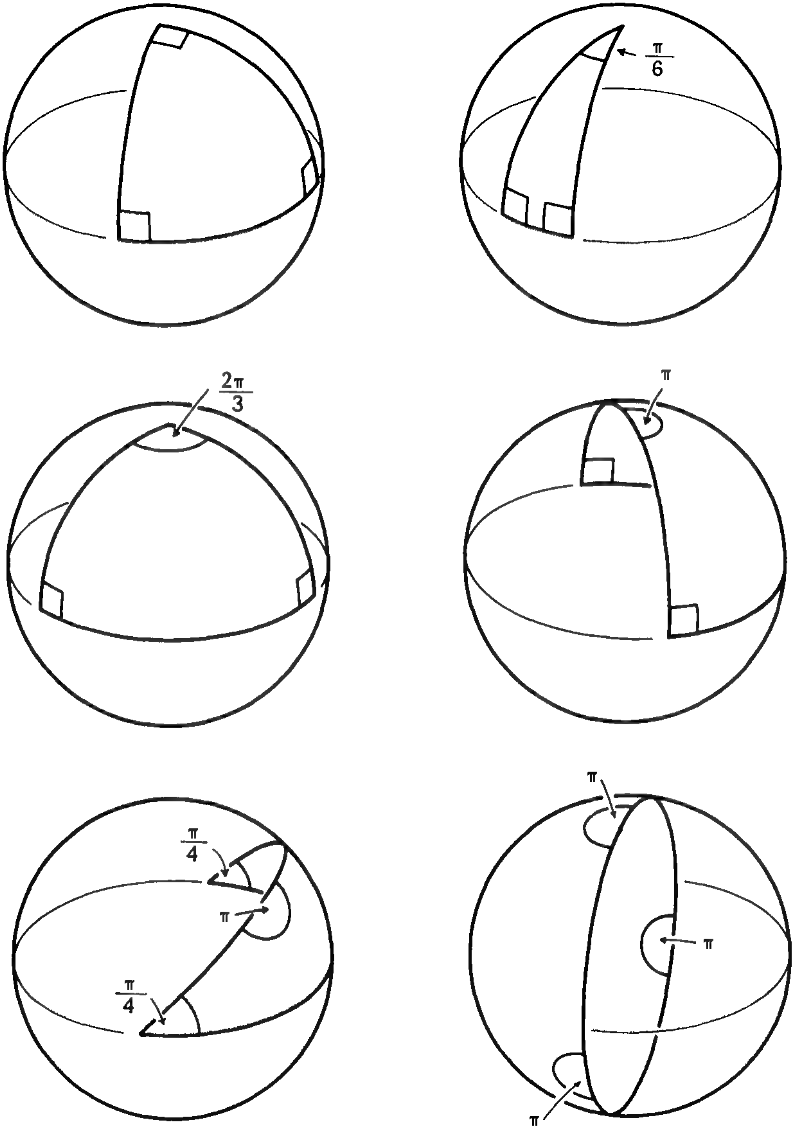


Figure 9.2 Some assorted spherical triangles. Three of the triangles are "degenerate" in the sense that each has one or more angles equal to π. The last triangle occupies an entire hemisphere, and its three sides all lie on the same great circle.

Even though there is no overwhelming need for a proof of the formula you just discovered, I would like to include one anyhow because it is so simple and elegant. (It is not, however, the sort of thing you're likely to stumble onto on your own. I struggled for hours without being able to prove the formula at all.)

First we have to know how to compute the area of a "double lune." A *double lune* is a region on a sphere bounded by two great circles, as shown in Figure 9.3. The largest the angle α can ever be is π, at which point the double lune fills up the entire sphere. So if α is, say, $\pi/3$, then we reason that since $\pi/3$ is $\frac{1}{3}$ the greatest possible angle π, the double lune must

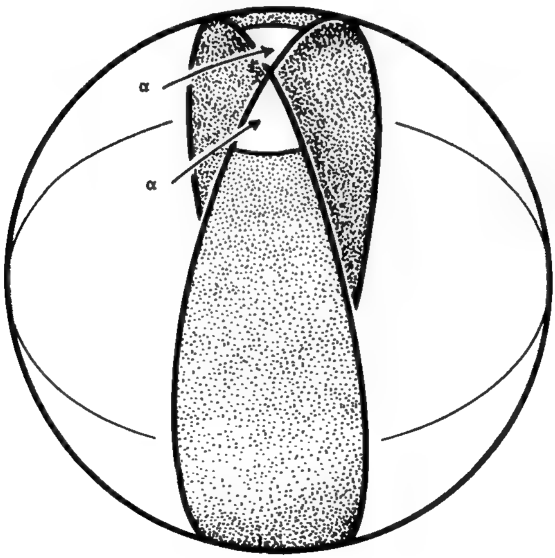


Figure 9.3 A double lune with angle α.

fill up ⅓ the area of the entire sphere, namely $(1/3)(4\pi) = 4\pi/3$. Using the same reasoning, we get that the area of a double lune with angle α is $(\alpha/\pi)(4\pi) = 4\alpha$. You can check this formula for some special cases, e.g. $\alpha = \pi/2$ or $\alpha = \pi$.

Now we'll find a formula for the area of a spherical triangle with angles α, β, and γ. First extend the sides of the triangle all the way around the sphere to form three great circles, as shown in Figure 9.4. An "antipodal triangle," identical to the original, is formed on the back side of the sphere. Figure 9.5

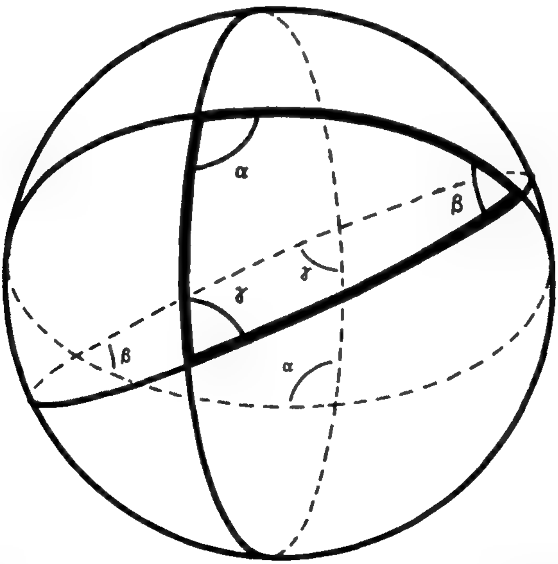


Figure 9.4 Extend the edges of the spherical triangle, and the resulting great circles will form an "antipodal triangle" on the back side of the sphere.

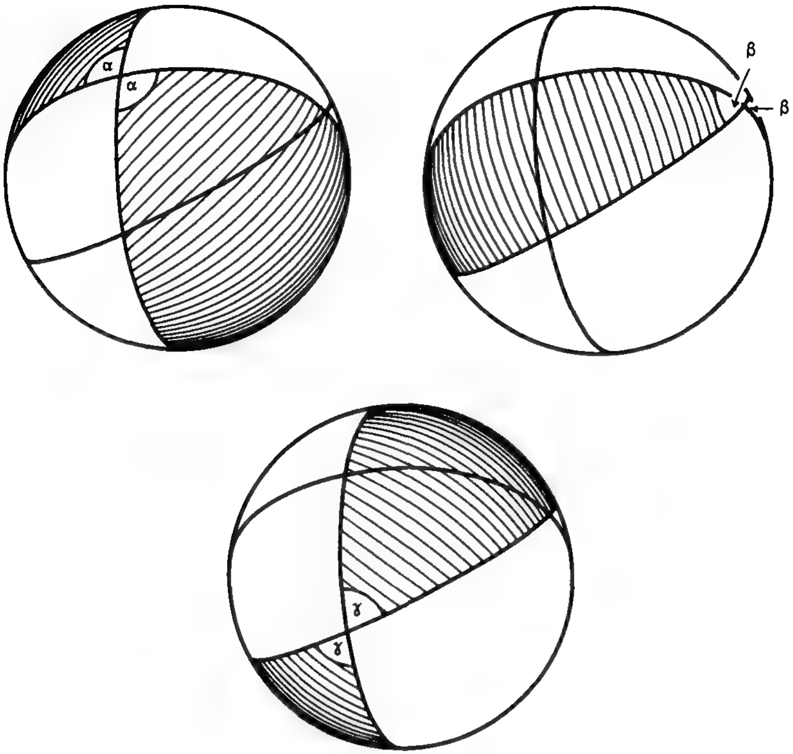


Figure 9.5 Three ways to shade in double lines.

shows three possible ways to shade in double lunes. These double lunes have respective angles α, β, and γ, and therefore their areas are 4α, 4β, and 4γ.

Now look what happens if we shade in all three double lunes simultaneously (Figure 9.6). All parts of the sphere get shaded in at least once, and the original and antipodal triangles each get shaded in three times (once for each double lune). So . . .

$$\begin{pmatrix}\text{area of}\\ \text{first}\\ \text{double}\\ \text{lune}\end{pmatrix} + \begin{pmatrix}\text{area of}\\ \text{second}\\ \text{double}\\ \text{lune}\end{pmatrix} + \begin{pmatrix}\text{area of}\\ \text{third}\\ \text{double}\\ \text{lune}\end{pmatrix} = \overset{\left[\begin{smallmatrix}\text{everything}\\ \text{was shaded}\\ \text{in once}\end{smallmatrix}\right]}{\begin{pmatrix}\text{area of}\\ \text{entire}\\ \text{sphere}\end{pmatrix}} \overbrace{+2\begin{pmatrix}\text{area of}\\ \text{original}\\ \text{triangle}\end{pmatrix} +2\begin{pmatrix}\text{area of}\\ \text{antipodal}\\ \text{triangle}\end{pmatrix}}^{\text{each triangle was shaded in two more times}}$$

$$4\alpha \quad + \quad 4\beta \quad + \quad 4\gamma \quad = \quad 4\pi \quad + \quad 2A \quad + \quad 2A$$

$$4(\alpha + \beta + \gamma) = 4(\pi + A)$$

$$\alpha + \beta + \gamma = \pi + A$$

$$(\alpha + \beta + \gamma) - \pi = A$$

which is just what we wanted to prove! In words, this formula says that the sum of the angles of a spherical triangle exceeds π by an amount equal to the triangle's area.

Exercise 9.3 The formula $(\alpha + \beta + \gamma) - \pi = A$ applies only to triangles on a sphere of radius one. How must you modify the formula to apply to triangles on a sphere of radius two? What about radius three? Write down a general formula for triangles on a sphere of radius r. □

Exercise 9.4 A society of Flatlanders lives on a sphere whose radius is exactly 1000 meters. A farmer has a triangular field with perfectly straight (i.e. geodesic) sides and angles which have been carefully measured as 43.624°, 85.123°, and 51.270°. What is the area of the field? Don't forget to convert the angles to radians. (Bonus Question: How accurately do you

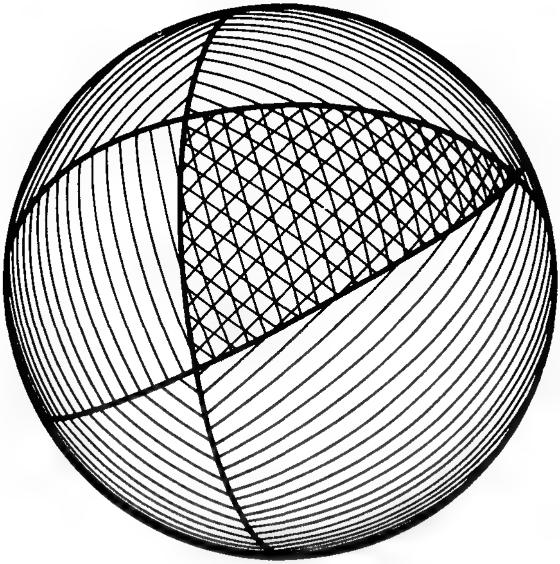


Figure 9.6 Look what happens when we shade in all three double lunes at once.

know the field's area? That is, by plus or minus what percent?) □

Exercise 9.5 A society of Flatlanders lives on a sphere. They carefully survey a triangle and find its angles to be 60.0013°, 60.0007° and 60.0011°, while its area is 5410.52 square meters. What is the area of the entire sphere? □

Exercise 9.6 Estimate the angle-sum of each of the following spherical triangles. Hint: First make a rough estimate of the area enclosed by each triangle, and then apply the formula from Exercise 9.3. Because you have to guess the areas of the triangles, you

will get only approximate, not exact, answers. The radius of the Earth is roughly 6400 km.

1. The triangle formed by Providence, Newport, and Westerly, Rhode Island. These cities are roughly 50 km apart.
2. The triangle formed by Houston, El Paso, and Amarillo, Texas. These cities are roughly 1000 km apart.
3. The triangle formed by Madras, India; Tokyo, Japan; and St. Petersburg, Russia. These cities are roughly 7000 km apart. □

We've now seen the first major way in which the geometry of a sphere differs from the geometry of a plane. Namely, the sum of the angles of a spherical triangle exceeds π by an amount proportional to the triangle's area, whereas the sum of the angles of a Euclidean (= flat) triangle equals π exactly (study Figure 9.7 for a proof of this last fact).

A piece of a sphere rips open when flattened onto a plane (Figure 9.8). This shows that a circle on a

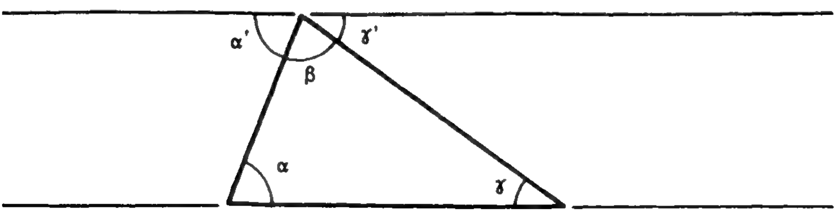


Figure 9.7 A quick proof that the sum of the angles in a Euclidean triangle is π: (1) $\alpha' + \beta + \gamma' = \pi$, (2) $\alpha = \alpha'$ and (3) $\gamma = \gamma'$, therefore (4) $\alpha + \beta + \gamma = \pi$.

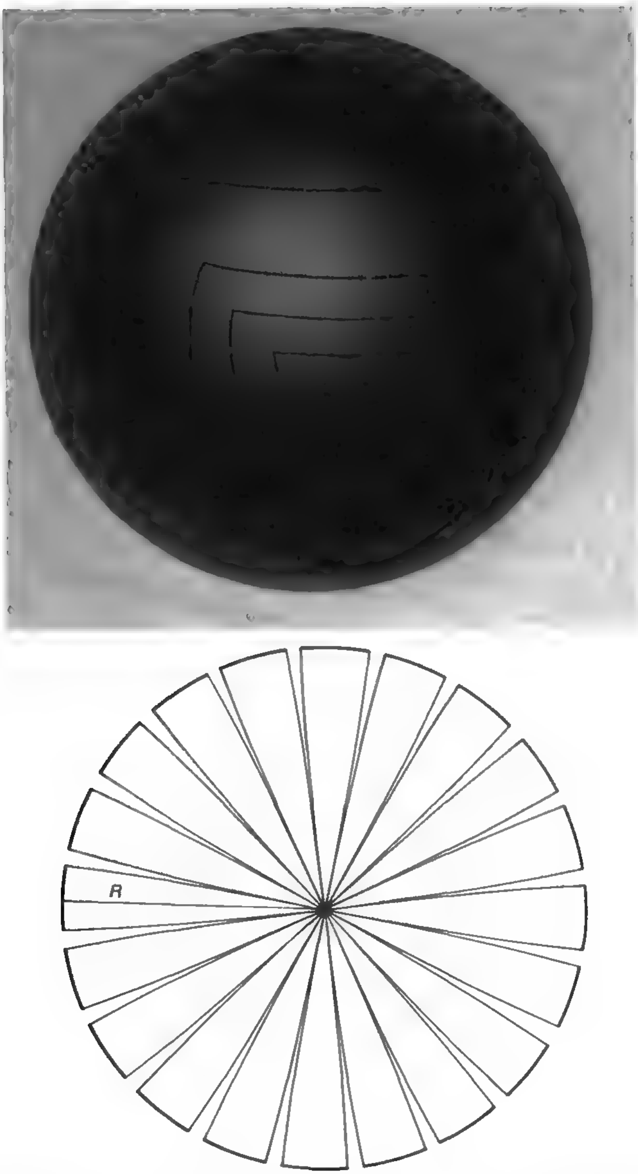


Figure 9.8 A piece of a sphere splits open when flattened.

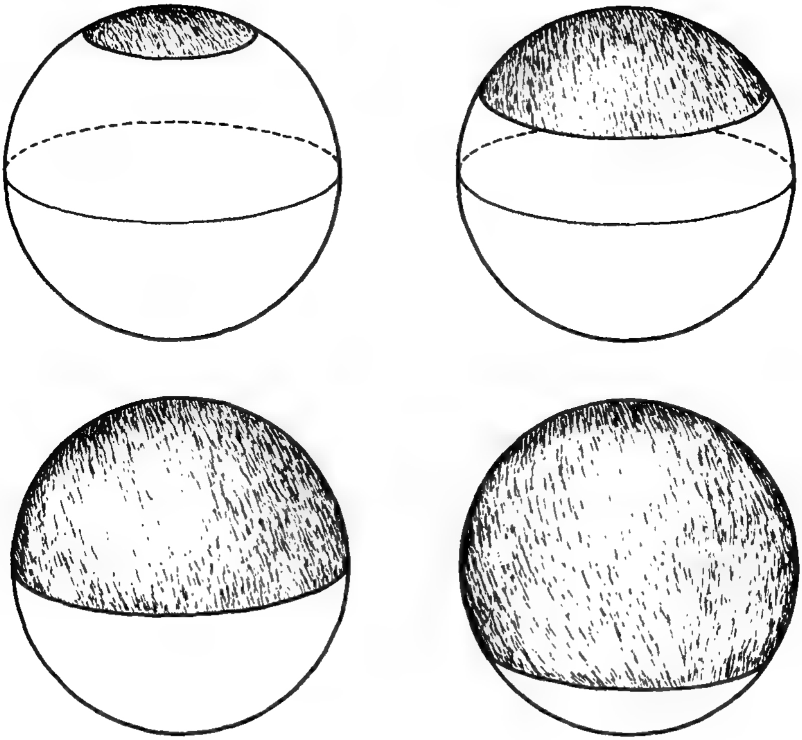


Figure 9.9 The circle's circumference first increases, but then decreases once the circle is past the equator.

sphere has a smaller circumference and encloses less area than a circle of the same radius in a plane. I should stress that the radius of a circle on a sphere is measured along the sphere itself—the way a Flatlander would measure it. Figure 9.9 shows that on a sphere a circle's circumference can actually shrink, even though the circle's (intrinsically measured) radius is still increasing.

Exercise 9.7 (*Review exercise*) What other surface has the local geometry of a sphere? Could Flatlanders living on this other surface tell that they weren't on a sphere? □

This Page Intentionally Left Blank

10

The Hyperbolic Plane

There are three types of homogeneous two-dimensional geometries, as illustrated in Figure 10.1. Two of these are already familiar. The first geometry shown is the familiar geometry of a sphere. This geometry is often called *elliptic geometry*, and is said to have *positive curvature*. The second geometry is the familiar geometry of the Euclidean plane. It is called *Euclidean geometry* and is said to have *zero curvature*. The third geometry shown is, loosely speaking, a saddle shaped geometry. It is less familiar than elliptic or Euclidean geometry, but certainly no less important.

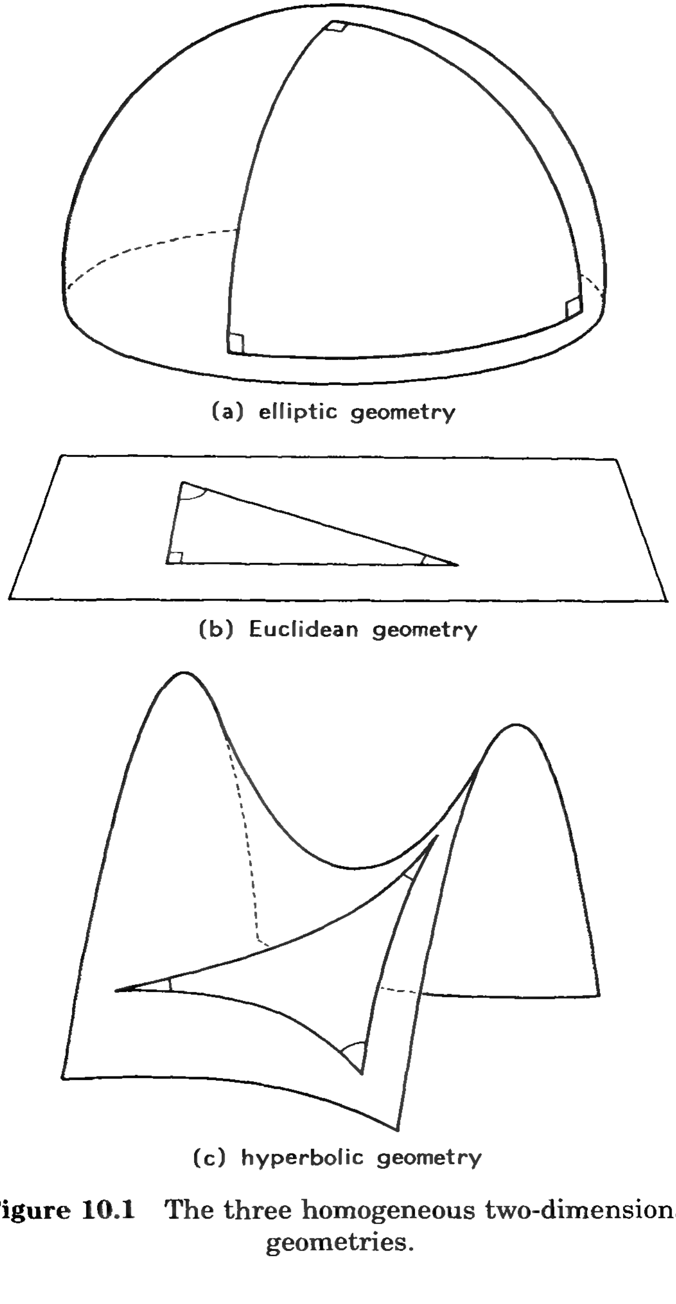
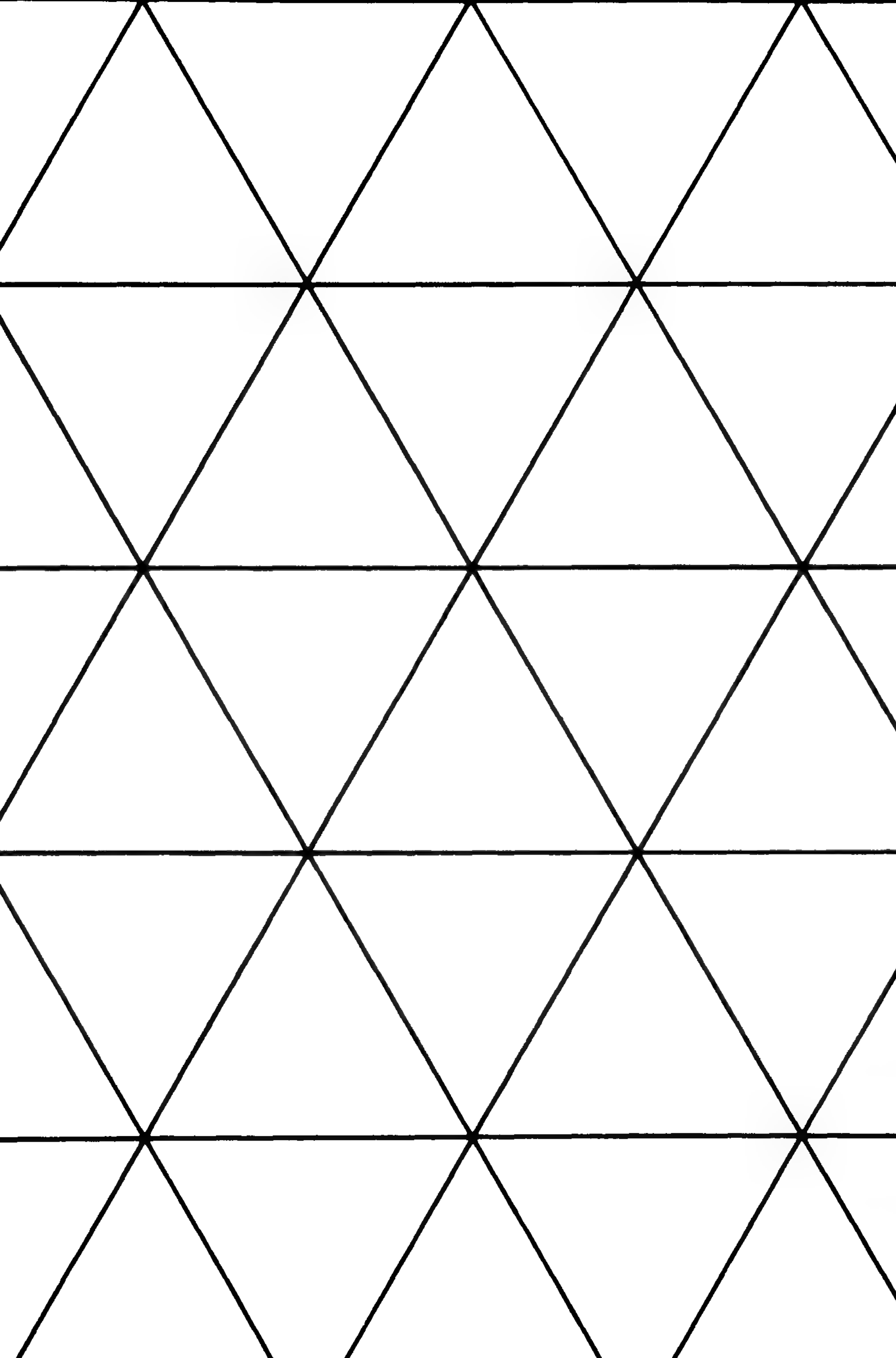


Figure 10.1 The three homogeneous two-dimensional geometries.

It is called *hyperbolic geometry* and is said to have *negative curvature*. The *hyperbolic plane* (abbreviated H^2) is an infinite plane that has hyperbolic geometry (= constant negative curvature) everywhere, just as the Euclidean plane is an infinite plane that has Euclidean geometry (= zero curvature) everywhere. It is a fact of nature, though, that there can be no infinite plane with elliptic geometry (= constant positive curvature) everywhere—it will, without fail, close back onto itself to form either a sphere or a projective plane.

While the saddle shape provides a good local picture of the hyperbolic plane, it fails to convey what the hyperbolic plane is like on a larger scale. To get a feel for the larger scale nature of the hyperbolic plane, make yourself some hyperbolic paper as described in Exercise 10.1. Handling the hyperbolic paper will do for your understanding what words and pictures cannot.

Exercise 10.1 *How to make hyperbolic paper.* (Thanks to Bill Thurston for this idea.) You'll need a large number of equilateral triangles, at least a hundred or so. The most efficient way to make them is to make several photocopies of Figure 10.2. Tape seven triangles together like the seven shaded triangles in Figure 10.3. Then add more triangles to the pattern in such a way that each vertex of the original figure is itself surrounded by seven triangles. Continue in this fashion as long as you want. The surface you get



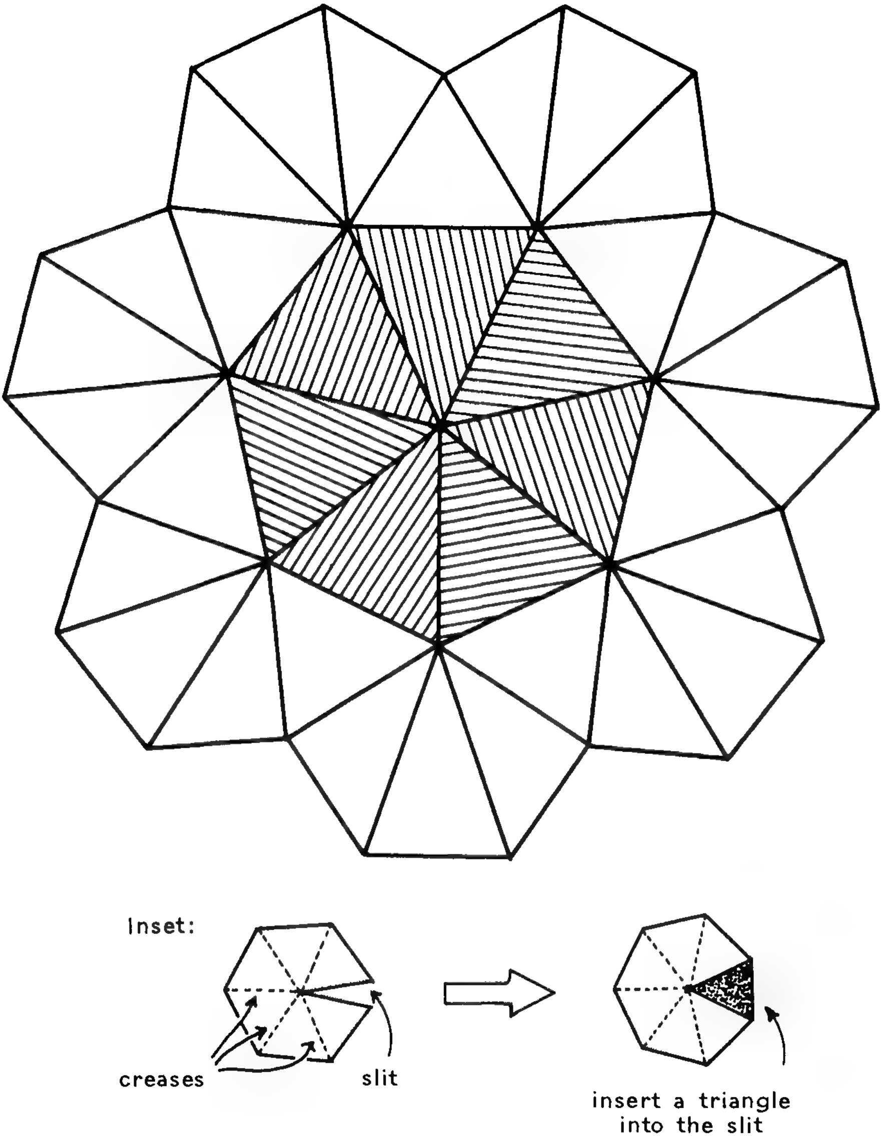


Figure 10.3 (*See Exercise 10.1*) Tape equilateral triangles together so that seven meet at each vertex. The inset shows a good way to get started.

will be very floppy, but this is how it should be. The larger the surface, the floppier and more accurate it will be. You can't go wrong so long as every vertex is surrounded by exactly seven triangles. A construction hint: Don't cut triangles apart only to tape them back together again! For example, rather than starting with seven separate triangles, you can start by inserting a single triangle into a block that has been slit and creased as shown in the inset to Figure 10.2. ◻

Exercise 10.2 What happens when you tape equilateral triangles together five per vertex instead of seven per vertex? Try it! ◻

Note that regions near the edge of your hyperbolic paper are indistinguishable from regions near the center. This demonstrates that the hyperbolic plane is homogeneous.

Elliptic and hyperbolic geometry tend to have contrasting properties; Euclidean geometry tends to serve as the borderline between them. Recall that a piece of a sphere rips open when flattened (Figure 9.8). In contrast, a piece of a hyperbolic plane wrinkles and overlaps when flattened (Figure 10.4).

More important for us will be the question of angles. In elliptic geometry the angles of a triangle add up to more than π radians (180°); on a sphere of radius one the angles exceed π by an amount exactly equal to the triangle's area [$(\alpha + \beta + \gamma) - \pi = A$]. In hyperbolic geometry the angles of a triangle add up to less than π radians; on the standard hyperbolic plane

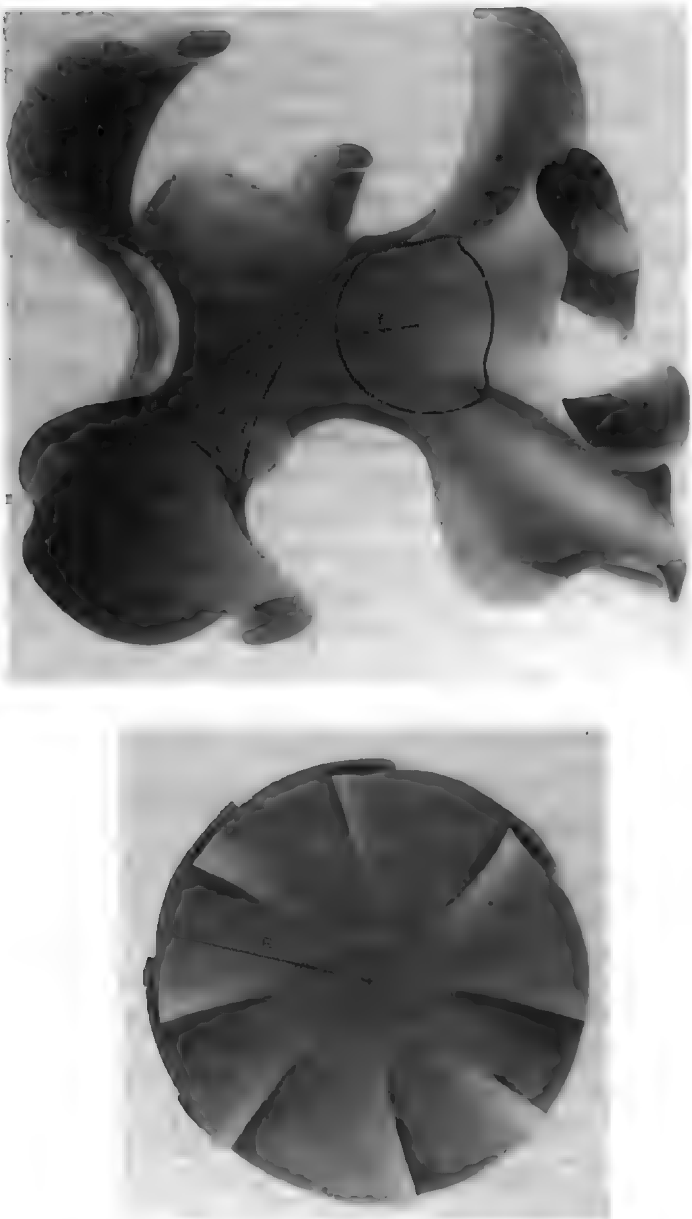


Figure 10.4 A piece of a hyperbolic plane wrinkles and overlaps when flattened.

the angles are less than π by an amount exactly equal to the triangle's area [$\pi - (\alpha + \beta + \gamma) = A$]. [But just as there can be smaller or larger spheres on which $(\alpha + \beta + \gamma) - \pi$ is proportional but not equal to the given triangle's area, so too can there be "larger" or "smaller" hyperbolic planes in which $\pi - (\alpha + \beta + \gamma)$ is proportional but not equal to the triangle's area. In this book, unless specified otherwise, "sphere" will refer to the sphere of radius one, for which $(\alpha + \beta + \gamma) - \pi = A$, and "hyperbolic plane" will refer to the standard hyperbolic plane for which $\pi - (\alpha + \beta + \gamma) = A$.] In the next chapter you will see why we are so interested in angles.

Exercise 10.3 In the hyperbolic plane, what is the area of a triangle whose angles are $\pi/3$, $\pi/4$, and $\pi/6$? □

It is a curious fact that no triangle in the hyperbolic plane can have an area greater than π. As the sides of a triangle get longer its angles get pointier, but even as the sides get infinitely long and the angles go to zero, the area never exceeds π. This is consistent with the formula $\pi - (\alpha + \beta + \gamma) = A$. I'll say more about this phenomenon in Chapter 15.

11

Geometries on Surfaces

The left side of Figure 11.1 shows the three surfaces you studied in Exercise 7.7. The first has cone points because the hexagon's corners are glued together in groups of only two. The second has a flat geometry with no cone points because the hexagon's corners are glued in groups of three. The third has the opposite of a cone point because all six corners of the hexagon are glued together. (You may want to review the material preceding Exercise 7.7. See in particular Figure 7.17.)

Both cone points and their opposites can be eliminated by utilizing the peculiar properties of the

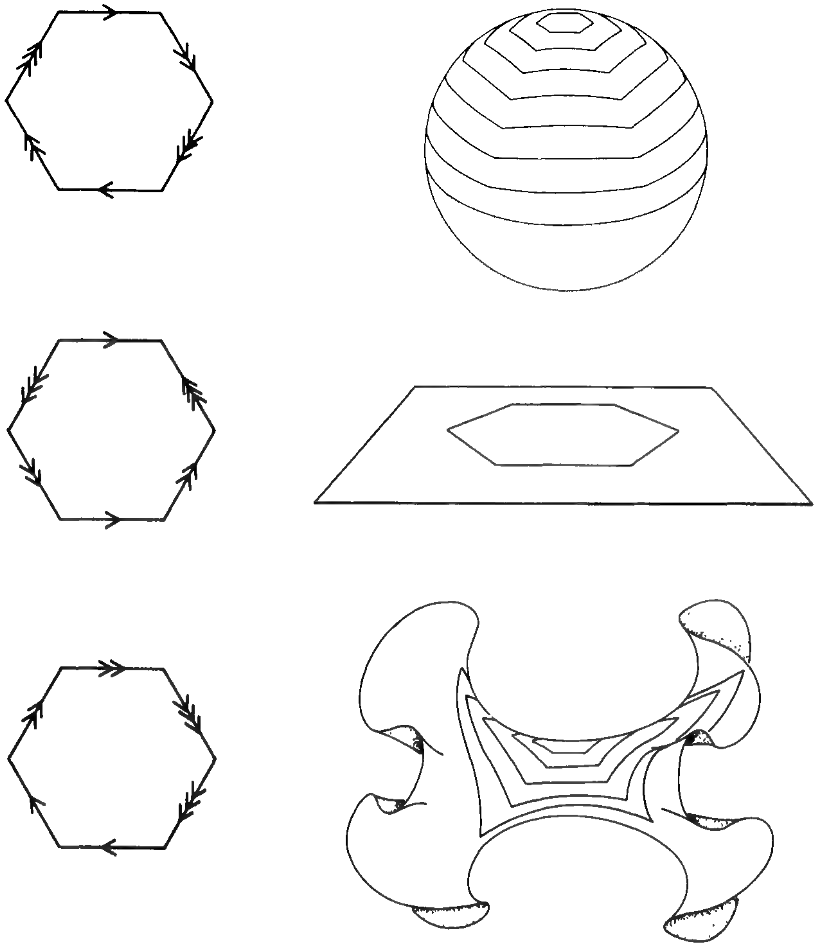


Figure 11.1 Every surface can be given a homogeneous geometry.

sphere and the hyperbolic plane. On a sphere, for example, larger hexagons have larger angles (just as larger triangles have larger angles). So to get rid of the cone points in the first example we need only put

the hexagon on a sphere and let it grow until its angles are big enough that the cone points disappear. This occurs when each angle is 180° and the hexagon fills an entire hemisphere. In the course of eliminating the cone points we have given the surface a homogeneous elliptic geometry. (By the way, do you recognize this surface?)

To get rid of the opposite of a cone point in the last example we have to put the hexagon on a hyperbolic plane, where larger polygons have smaller angles. If we let the hexagon grow until each angle shrinks to 60°, then the opposite-of-a-cone-point will disappear. In the process the surface will acquire a homogeneous hyperbolic geometry.

Exercise 11.1 The surface in Figure 7.16 has two cone points. How can they be eliminated? What homogeneous geometry does this surface acquire? By the way, what surface is this? □

Exercise 11.2 For each surface in Figure 11.2 use the "walking around corners" technique of Chapter 7 to determine how the polygon's corners fit together and whether the surface has cone points. Which of the surfaces can be given elliptic geometry, which can be given Euclidean geometry, and which can be given hyperbolic geometry? □

It turns out that *every* surface can be given some homogeneous geometry. You can modify the above technique to find homogeneous geometries for sur-

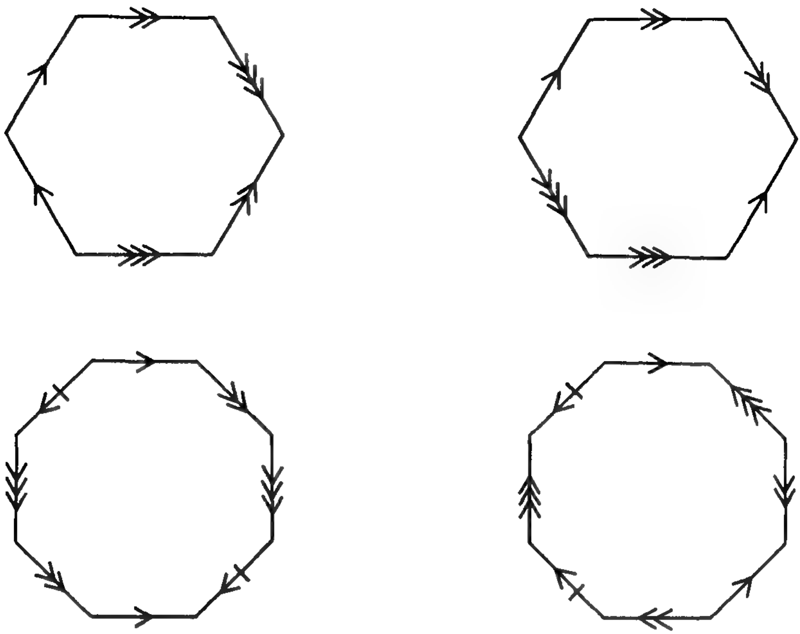


Figure 11.2 Find a homogeneous geometry for each of these surfaces.

faces that are drawn in three dimensions rather than as glued polygons. For example, say you want to find a homogeneous geometry for a three-holed doughnut surface. First cut the surface into hexagons whose corners meet in groups of four (Figure 11.3). Deform each hexagon to be regular (like the hexagons on the left in Figure 11.1). Regular hexagons have 120° corner angles—which are too big to fit together in groups of four—so put each hexagon in the hyperbolic plane and let it expand until its angles decrease to 90°. Each hexagon now has both (1) a homogeneous hyperbolic

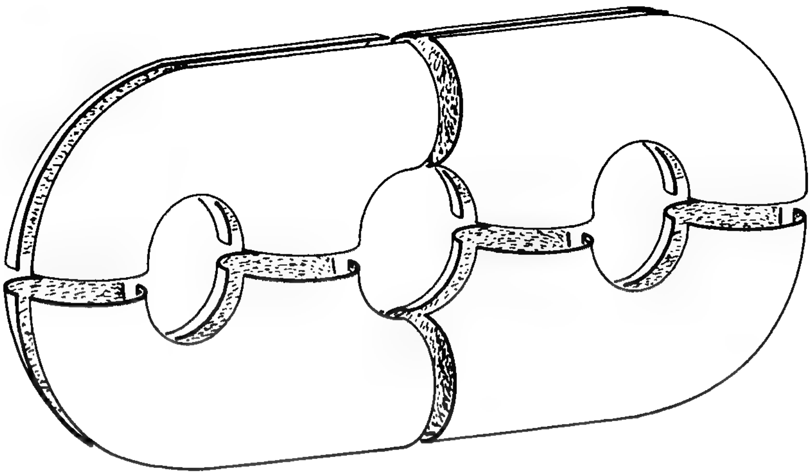


Figure 11.3 A three-holed doughnut surface can be divided into eight hexagons whose corners meet in groups of four.

geometry, and (2) corner angles of 90°. Now reglue the hexagons to reassemble the surface. Don't imagine physically regluing them in three-dimensional space —this can't be done without deforming them and ruining their hyperbolic geometry. Instead imagine each to be reglued to its previous neighbors in the abstract sense that a Flatlander leaving one hexagon reappears in a neighboring one. This abstractly reassembled surface has the same global topology as the original three-holed doughnut. But unlike the original three-holed doughnut, whose local geometry varied irregularly from point to point, this new surface has the homogeneous local geometry of the hyperbolic plane. The hexagons' 90° corners fit perfectly in groups of four, so there are no cone points.

Exercise 11.3 Draw a picture showing how to cut a four-holed doughnut surface (T^2 # T^2 # T^2 # T^2) into hexagons whose corners meet in groups of four. Do the same for a two-holed doughnut surface (T^2 # T^2). How many hexagons do you get when you cut up an n-holed doughnut surface? □

Any doughnut surface with at least two holes can be cut into hexagons whose corners meet in groups of four. Any such surface can, therefore, be given a hyperbolic geometry by the above technique. A one-holed doughnut surface cannot be given a hyperbolic geometry, but it can be given a Euclidean geometry (flat torus). A zero-holed doughnut surface, i.e. a sphere, also cannot be given a hyperbolic geometry, but its usual elliptic geometry is homogeneous. As stated above, every surface can be given some type of homogeneous geometry. In Chapter 12 we'll see that no surface can be given more than one type.

Exercise 11.4 The connected sum of two projective planes can be cut into two squares (Figure 11.4). Similarly, P^2 # P^2 # P^2 can be cut into two hexagons, P^2 # P^2 # P^2 # P^2 can be cut into two octagons, etc. In each case the polygon's corners meet in groups of four (two corners on one side of a rim are glued to the two corners on the opposite side). Which of these surfaces can be given which homogeneous geometry? □

Exercise 11.5 Recall that every surface is topologically equivalent to a surface on the following list:

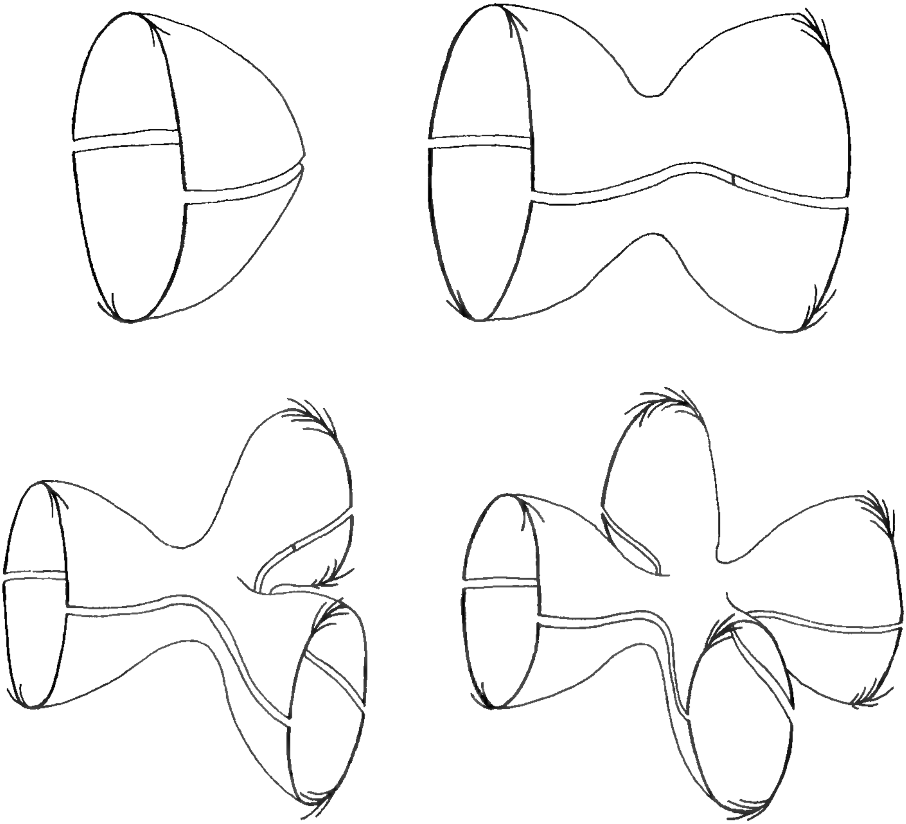


Figure 11.4 The connected sum of n projective planes can be divided into two 2n-gons whose corners meet in groups of four.

	T^2	$T^2 \# T^2$	$T^2 \# T^2 \# T^2$	$T^2 \# T^2 \# T^2 \# T^2$	...
S^2					
	P^2	$P^2 \# P^2$	$P^2 \# P^2 \# P^2$	$P^2 \# P^2 \# P^2 \# P^2$	...

Place each surface on the list into an appropriate box in the table of Figure 11.5 according to its orientability and to which homogeneous geometry it can be given. □

	orientable	nonorientable
elliptic		
Euclidean		
hyperbolic		

Figure 11.5 Surfaces may be grouped according to their geometry and orientability.

Except for the sphere and the projective plane, which have elliptic geometry, and the torus and the Klein bottle, which have Euclidean (flat) geometry, *all* surfaces can be given hyperbolic geometry. When a surface can be given a certain homogeneous geometry, one says that the surface "admits" that geometry. Thus the sphere and the projective plane admit elliptic geometry, the torus and the Klein bottle admit Euclidean geometry, and all other surfaces admit hyperbolic geometry.

12

The Gauss–Bonnet Formula and the Euler Number

Every surface has an *Euler number*, an integer that contains essential information about the surface's global topology. ("Euler" is pronounced "oiler.") The Euler number is easy to compute, and it immediately predicts which homogeneous geometry the surface will admit: surfaces with positive Euler number admit elliptic geometry, surfaces with zero Euler number admit Euclidean geometry, and surfaces with negative Euler number admit hyperbolic geometry. In fact, the Euler number is so powerful that if you know a surface's Euler number and you know whether it's or-

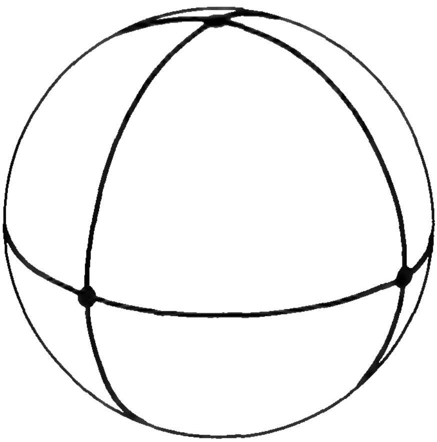
ientable or not, then you can immediately say what global topology the surface has! The Gauss–Bonnet formula relates a surface's Euler number to its area and curvature. ("Bonnet" is a French name, so the "t" is silent and the stress is on the second syllable, "buh-NAY.")

The Euler number is defined in terms of something called a cell-division, so we'll start by saying just what *cells* and *cell-divisions* are. A zero-dimensional cell is a point (usually called a vertex). A one-dimensional cell is topologically a line segment (usually called an edge). And a two-dimensional cell is topologically a polygon (usually called a face, even if it isn't the face *of* anything). A cell-division is what you get when you divide a surface into cells. It's similar to the decomposition of a surface into polygons used in Chapter 11 (see Figure 11.3), only now we're interested in the vertices and edges as well as the faces.

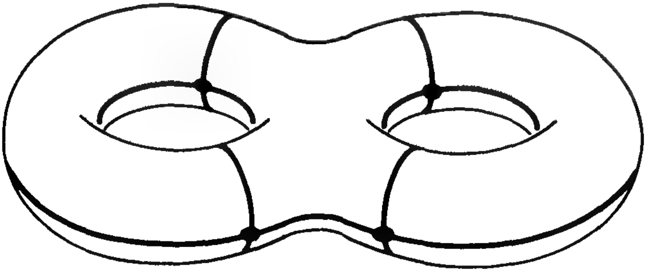
Figure 12.1 shows some typical cell-divisions. The first is a cell-division of S^2 consisting of six vertices, twelve edges, and eight triangular faces. The second is a cell-division of $T^2 \# T^2$ consisting of eight vertices, sixteen edges, and six faces (the two faces in the middle—one on top and one on bottom—are topologically octagons, while the two on each end are topologically squares). The third is a cell-division of T^2. It contains *one* vertex—not four—because the four corners of the square meet at a single point (see inset, and recall Figure 3.3). Similarly, it contains only two edges, because the square's four edges are glued together in

pairs. Finally, this cell-division obviously contains only one face. The last drawing in Figure 12.1 shows a cell-division of $\mathbb{P}^2$.

Exercise 12.1 Count the number of vertices, edges and faces in the cell-division of $\mathbb{P}^2$ given in Figure 12.1(d). Be careful: vertices and edges on the disk's circumference are glued together in pairs. □

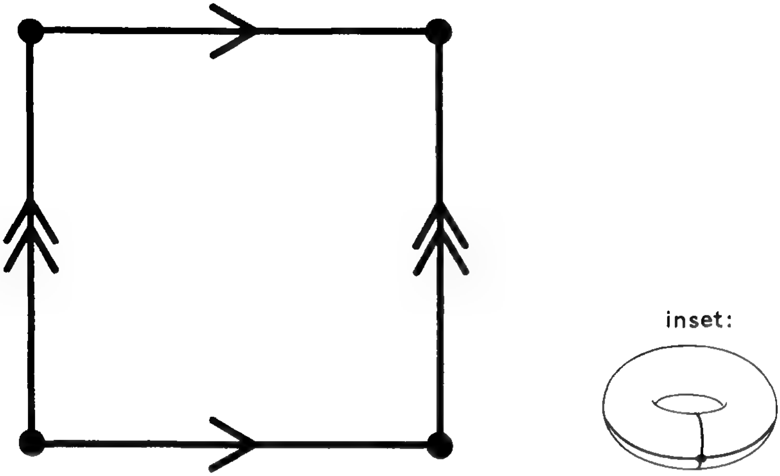


(a) a cell-division of S^2

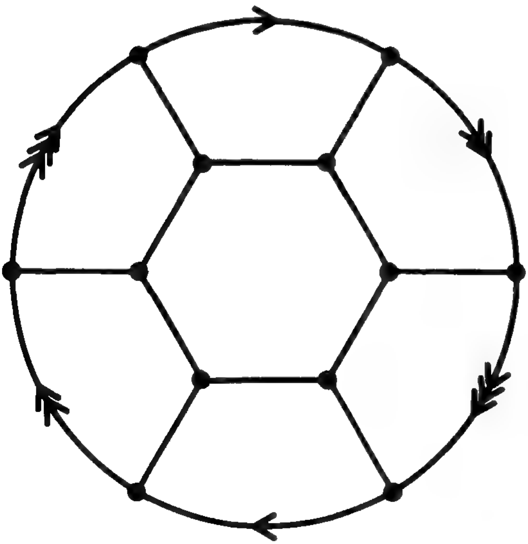


(b) a cell-division of $T^2 \# T^2$

Figure 12.1 Some typical cell-divisions.



(c) a cell-division of T^2



(d) a cell-division of P^2

Figure 12.1 Continued.

From now on

v = the number of vertices in a cell-division
e = the number of edges in a cell-division
f = the number of faces in a cell-division

So, for example, the above cell-division of S^2 has $v = 6$, $e = 12$ and $f = 8$, while the cell-division for $T^2 \# T^2$ has $v = 8$, $e = 16$ and $f = 6$.

By the way, we will usually assume that the edges in a cell-division are geodesics.

To "discover" the Gauss–Bonnet formula and the Euler number we'll need a couple facts about polygons. A polygon with n sides is called an *n-gon*.

Exercise 12.2 In the Euclidean plane the angles of any triangle add up to exactly π radians. Use Figure 12.2 (ignore the formulas) to deduce that the angles of a Euclidean n-gon add up to $(n - 2)\pi$ radians. Thus the angles of a quadrilateral add up to 2π, the angles of a pentagon add up to 3π, etc. □

Exercise 12.3 The area of a triangle on a unit sphere is the sum of its angles minus π [in symbols $A = (\alpha + \beta + \gamma) - \pi$]. Deduce that the area of an n-gon on a unit sphere is A = (sum of all angles) $- (n - 2)\pi$. (Hint: The area of an n-gon is the sum of the areas of $n - 2$ triangles, as per Figure 12.2.) By referring to the previous exercise you can interpret this formula as saying that the area of a spherical n-gon equals the

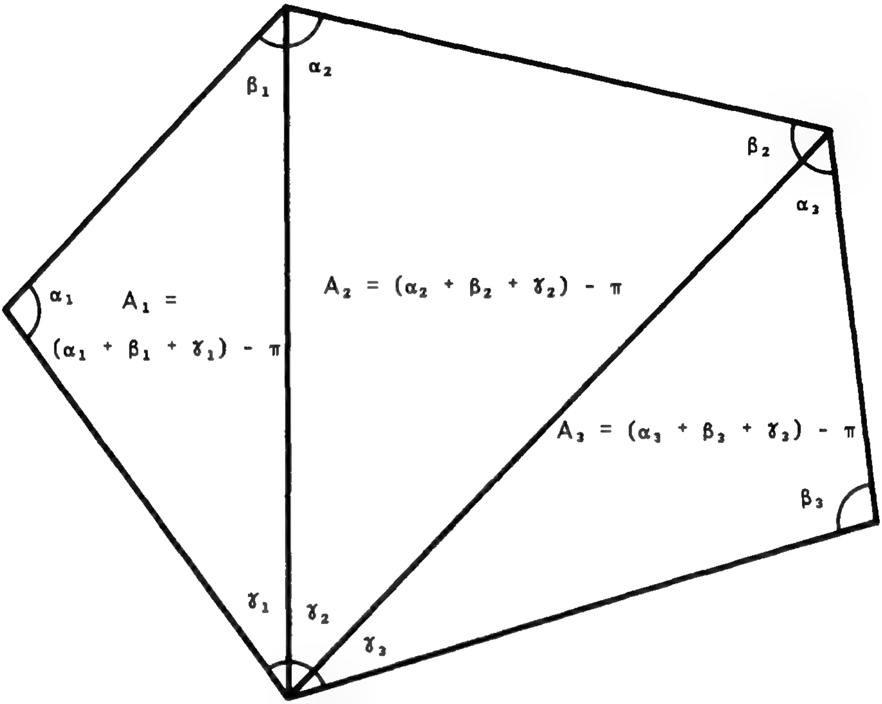


Figure 12.2 An n-gon can be divided into $n - 2$ triangles.

difference between what the angles actually are and what they would have been if the n-gon were flat. □

Exercise 12.4 Find a formula for the area of an n-gon in the hyperbolic plane. □

It is now fairly easy to "discover" the Gauss–Bonnet formula and the Euler number. We'll do it for a sphere first. Give the sphere a cell-division—any cell-division will do. The plan is to express the area of the sphere

as the sum of the areas of the faces of the cell-division, and then use the formula A = (sum of angles) − $(n - 2)\pi$ to get a new expression for the area of each face. Then we'll follow our noses and see what happens.

I'll first write out the computation in its entirety, then I'll go back and explain each of the steps. (For an extra challenge, you might want to see how much of the computation you can figure out on your own *before* you read the explanations.)

$$
\begin{aligned}
&(1)\ \begin{matrix}\text{Total}\\ \text{Area}\end{matrix} = \left[\begin{matrix}\text{area of}\\ \text{first face}\end{matrix}\right] + \left[\begin{matrix}\text{area of}\\ \text{second face}\end{matrix}\right] + \cdots + \left[\begin{matrix}\text{area of}\\ \text{last face}\end{matrix}\right]\\
&(2)\qquad = \left[\left(\begin{matrix}\text{sum of angles}\\ \text{of first face}\end{matrix}\right) - (n_1 - 2)\pi\right] + \cdots + \left[\left(\begin{matrix}\text{sum of angles}\\ \text{of last face}\end{matrix}\right) - (n_f - 2)\pi\right]\\
&(3)\qquad = \left[\left(\begin{matrix}\text{sum of angles}\\ \text{of first face}\end{matrix}\right) + \cdots + \left(\begin{matrix}\text{sum of angles}\\ \text{of last face}\end{matrix}\right)\right] - (n_1 + n_2 + \cdots + n_f)\pi + (2 + 2 + \cdots + 2)\pi\\
&(4)\qquad = \left[\begin{matrix}\text{sum of all angles in}\\ \text{the entire cell-division}\end{matrix}\right] - (n_1 + n_2 + \cdots + n_f)\pi + (2 + 2 + \cdots + 2)\pi\\
&(5)\qquad = 2\pi v - 2\pi e + 2\pi f\\
&(6)\qquad = 2\pi(v - e + f)
\end{aligned}
$$

Step-by-Step Explanation

Step 1 Write the area of the surface as the sum of the areas of the faces of the cell-division.

Step 2 Use the formula A = (sum of angles) − $(n - 2)\pi$ (from Exercise 12.3) to get a new expression for the area of each face.

Step 3 Regroup terms.

Step 4 Note that adding up all the angles of all the faces is the same as adding up all the angles in the entire cell-division.

Step 5 There are three things going on at this step. ***Substep A*** The sum of all the angles in a cell-division is $2\pi v$. This is because there are v vertices, and the sum of the angles surrounding each vertex is 2π. For a specific example, consider the cell-division of S^2 shown in Figure 12.1(a). We could have computed the sum of all its angles as

$$\begin{aligned}&(\text{sum of angles in each face}) \times (\text{number of faces})\\ &\quad = (\pi/2 + \pi/2 + \pi/2) \times 8 = (3\pi/2) \times 8 = 12\pi\end{aligned}$$

but instead we compute it as

$$\begin{aligned}&(\text{sum of angles surrounding each vertex})\\ &\times (\text{number of vertices}) = 2\pi \times v = 2\pi \times 6 = 12\pi\end{aligned}$$

Substep B Note that $n_1 + n_2 + \cdots + n_f = 2e$, where e is the total number of edges in the cell-division. Each n represents the number of edges of a certain face. By adding up all the n's we are adding up all the edges in the entire cell-division. However, each edge is the boundary of two faces, so each edge gets counted twice. This is why $n_1 + n_2 + \cdots + n_f$ equals $2e$ rather than just e. A specific example might clarify things. The cell-division of S^2 shown in Figure 12.1(a) has 8 faces, each of which has 3 edges. Therefore it might seem that the cell-division has $8 \times 3 = 24$ edges. But each edge was counted twice—really there are only 12 edges.

Substep C The last part of Step 5 is the substitution $2 + 2 + \cdots + 2 = 2f$, where f is the number of faces in the cell-division. This is easy to understand: back in line (1) there was one term for each face in the cell-division, so now the sum $2 + 2 + \cdots + 2$ must contain one 2 for each face.

Step 6 Factor out 2π.

The formula Area $= 2\pi(v - e + f)$ is a stunning conclusion! We know that a unit sphere has area 4π, so the formula says that $v - e + f$ must be two, *no matter what cell-division we choose!*

Exercise 12.5 Check that $v - e + f = 2$ for the cell-division of Figure 12.1(a). Then draw any other cell-division of S^2 (regular or irregular) and check that $v - e + f = 2$ for it as well. □

The derivation of the formula $A = 2\pi(v - e + f)$ did not rely on the fact that the surface in question was a sphere. It required only that the surface have the *local* geometry of a unit sphere. Therefore the computations apply equally well to a projective plane of radius one. The projective plane is made from a hemisphere, which has only half the area of a sphere, so in this case the formula $A = 2\pi(v - e + f)$ becomes $2\pi = 2\pi(v - e + f)$ and we conclude that $v - e + f$ must equal one for *every* cell-division of P^2!

Exercise 12.6 Check that $v - e + f = 1$ for the cell-division of P^2 shown in Figure 12.1(d) and referred to in Exercise 12.1. Then draw some other cell-division of P^2 and check that $v - e + f = 1$ for it too. □

The formula $A = 2\pi(v - e + f)$ is the *Gauss–Bonnet formula* for surfaces with elliptic geometry. The quantity $v - e + f$ is the *Euler number*; it is invariably denoted by the Greek letter χ (spelled "chi" and pronounced "kai"). Thus $\chi \equiv v - e + f$, and the Gauss–Bonnet formula for elliptic geometry can be written $A = 2\pi\chi$. Because $\chi \equiv v - e + f = 2$ for every cell-division of S^2, we say that "the Euler number of a sphere is two," and write $\chi(S^2) = 2$. Similarly, a projective plane has Euler number one and we write $\chi(P^2) = 1$.

Exercise 12.7 Modify the derivation of $A = 2\pi\chi$ to discover an analogous Gauss–Bonnet formula for surfaces with hyperbolic geometry. □

Exercise 12.8 Use the cell-division of Figure 12.1(b) to compute $\chi(T^2 \# T^2)$. To check your answer, recompute $\chi(T^2 \# T^2)$ using a different cell-division (for instance, you could divide $T^2 \# T^2$ into four hexagons as in Exercise 11.3). □

Exercise 12.9 What is the area of a two-holed doughnut with (standard) hyperbolic geometry? Hint: Combine your answers from the preceding two exercises. □

Exercise 12.10 Compute the Euler number of a three-holed doughnut. What is the area of a three-holed doughnut with hyperbolic geometry? □

There is also a Gauss–Bonnet formula for surfaces with Euclidean geometry. In Euclidean geometry the angles of an n-gon add up to exactly $(n-2)\pi$ (recall Exercise 12.2), so for any cell division

$$\begin{pmatrix}\text{sum of angles}\\ \text{of first face}\end{pmatrix} + \cdots + \begin{pmatrix}\text{sum of angles}\\ \text{of last face}\end{pmatrix} = (n_1 - 2)\pi + \cdots + (n_k - 2)\pi$$

$$\begin{pmatrix}\text{sum of all angles in}\\ \text{the entire cell division}\end{pmatrix} = (n_1 + n_2 + \cdots + n_k)\pi - (2 + 2 + \cdots + 2)\pi$$

$$2\pi v \qquad = \qquad 2\pi e \qquad - \qquad 2\pi f$$

$$2\pi(v - e + f) = 0$$

where the steps follow as in the earlier derivation of the elliptic Gauss–Bonnet formula. Since $0 = 2\pi(v - e + f) = 2\pi\chi$, it follows that a surface with Euclidean geometry must have Euler number zero!

Exercise 12.11 Compute the Euler number of a torus using the cell-division of Figure 12.1(c). Then recompute it using a different cell-division of your own design. Compute $\chi(K^2)$ using two different cell-divisions. □

The three Gauss–Bonnet formulas can be summarized as $kA = 2\pi\chi$, where $k = -1, 0,$ or $+1$ according to whether the surface has hyperbolic geometry (*negative* curvature), Euclidean geometry (*zero* curvature) or elliptic geometry (*positive* curvature). Later in the chapter we'll see more general forms of the Gauss–

Bonnet formula, but even now we can deduce that no surface admits more than one homogeneous geometry: the "curvature" of the geometry ($k = -1, 0, +1$) must have the same sign as the Euler number. Intuitively, the surface's global topology determines which homogeneous local geometry "fits" it.

Exercise 12.12 In previous exercises you have computed $\chi(S^2)$, $\chi(T^2)$, $\chi(T^2 \# T^2)$ and $\chi(T^2 \# T^2 \# T^2)$. What is the pattern? What do you think $\chi(T^2 \# T^2 \# T^2 \# T^2)$ is? □

It's easy to derive a general formula for the Euler number of the connected sum of n tori. In Exercise 11.3 you found that an n-holed doughnut can be divided into $4n - 4$ hexagons whose corners meet in groups of four. Once we figure out how many edges and vertices such a cell-division has it will be easy to compute the Euler number.

1. There are $6 \times (4n - 4) \div 2 = 12n - 12$ edges. *Reason*: Each of the $4n - 4$ hexagons contributes six edges, for a total of $24n - 24$, but we have to divide by two to compensate for the fact that each edge has been counted twice because it's the border of two hexagons.

2. There are $6 \times (4n - 4) \div 4 = 6n - 6$ vertices. *Reason*: Each of the $4n - 4$ hexagons contributes six vertices, for a total of $24n - 24$, but we have to divide by four to compensate for the fact that each vertex has been counted four times because it's the corner of four hexagons.

The Euler number of the connected sum of n tori is $\chi = v - e + f = (6n - 6) - (12n - 12) + (4n - 4) = 2 - 2n$ in agreement with the pattern established in Exercise 12.12.

Exercise 12.13 The connected sum of n projective planes can be divided into two $2n$-gons whose corners meet in groups of four (Figure 11.4). Compute the Euler number. Complete the table in Figure 12.3. □

Euler number	*orientability*: orientable	nonorientable
2	S^2	
1		P^2
0	T^2	
-1		
-2	$T^2\#T^2$	
-3		
-4	$T^2\#T^2\#T^2$	
-5		
-6	$T^2\#T^2\#T^2\#T^2$	
-7		
-8	$T^2\#T^2\#T^2\#T^2\#T^2$	
.		
.	etc.	
.		

Figure 12.3 Euler number and orientability completely determine a surface's topology.

A priori it could be very difficult to decide whether two manifolds are topologically the same. In fact, the problem of deciding whether two *three-dimensional* manifolds are topologically the same is so difficult that in spite of decades of work by many people, no one has yet found a foolproof procedure to do it. (There are many *practical* procedures to tell two three-manifolds apart, but each of these procedures can be "fooled" by some appropriate pair of similar-looking manifolds.) In this light it is amazing that, according to the table in Figure 12.3, we can conclusively identify a surface just by computing its Euler number and deciding whether it's orientable or not! For example, say a certain surface has Euler number -4 and is orientable, then by consulting the table we can conclude that the surface must be topologically equivalent to T^2 # T^2 # T^2. Similarly, a nonorientable surface with Euler number -2 must be P^2 # P^2 # P^2 # P^2.

Exercise 12.14 For each surface in Figure 11.1, compute its Euler number (be careful counting those edges and vertices!), determine its orientability, and use the table in Figure 12.3 to identify the surface. Do the same for the surfaces in Figure 11.2. □

We'd like to apply the Gauss–Bonnet formula $kA = 2\pi\chi$ to all homogeneous surfaces, not just those with "standard" elliptic geometry (the geometry of a sphere of radius one), "standard" hyperbolic geometry, or Euclidean geometry. Consider, for example, how we might apply the formula to a sphere of radius three.

A sphere of radius three has the same Euler number as a sphere of radius one, but its area is nine times as great. Therefore its "curvature" k must be only one-ninth as great if the formula $kA = 2\pi\chi$ is still to hold. This is not unreasonable: larger spheres certainly *look* less curved than smaller spheres (Figure 12.4). In general, we define the curvature of a sphere of radius r to be $k \equiv 1/r^2$. A projective plane with the local geometry of a sphere of radius r also has curvature $k \equiv 1/r^2$.

Exercise 12.15 If distance is measured in meters, in what units is curvature measured? Do the units of kA match the units of $2\pi\chi$? □

Exercise 12.16 Compute the curvature, area, and Euler number of a projective plane of radius two meters, and check that the Gauss–Bonnet formula holds. (The area of a *sphere* of radius r is $4\pi r^2$.) □

We can rephrase the above definition of curvature by saying that if we start with a sphere of radius one

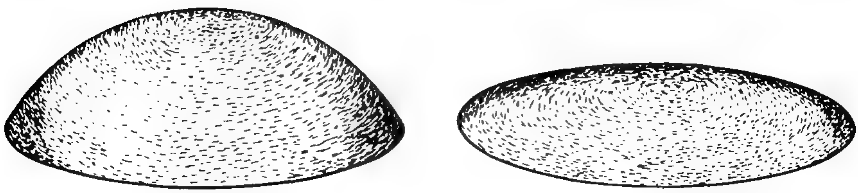


Figure 12.4 A piece of a small sphere is more curved than a piece of a large sphere.

and expand all distances by a factor of r ($r > 1$ means the sphere actually expands, $r < 1$ means it shrinks) then the resulting sphere will have curvature $k \equiv 1/r^2$. This rephrased definition is awkward, but, unlike the first definition, it avoids the concept of "radius" and can be adapted to the hyperbolic plane. The standard hyperbolic plane has curvature $k \equiv -1$. If we expand all distances in the standard hyperbolic plane by a factor of r, then the curvature of the resulting hyperbolic plane is defined to be $k \equiv -1/r^2$. Any surface with the local geometry of such a hyperbolic plane also has curvature $k \equiv -1/r^2$. For example, if $T^2 \# T^2$ is given a standard hyperbolic geometry, its curvature is -1, its area is 4π, and the Gauss–Bonnet formula reads $(-1)(4\pi) = 2\pi(-2)$. But when the surface is enlarged by a factor of five, then the curvature becomes $-1/25$, the area becomes $25 \times 4\pi = 100\pi$, and the Gauss–Bonnet formula reads $(-1/25)(100\pi) = 2\pi(-2)$ (correct again!).

Exercise 12.17 A society of Flatlanders lives in a universe with the global topology of $T^2 \# T^2 \# T^2$ and a homogeneous local geometry of constant curvature -0.00001 (meters)$^{-2}$. What is the area of their universe? □

We can refer to any surface with a homogeneous geometry as a "surface of constant curvature k," where the exact value of k depends on the geometry of the surface in question. We're talking about a surface with elliptic geometry when k is positive, with Euclidean

geometry when k is zero, and with hyperbolic geometry when k is negative. The closer k is to zero, the flatter the surface is, and the further k is from zero (in either direction), the more curved (either positively or negatively) the surface is. The curvature k is usually called the *Gaussian curvature.* We can use this terminology to summarize our results as

> **The Unified Gauss–Bonnet Formula for Surfaces of Constant Curvature**
>
> *If a surface has area A, Euler number χ, and constant Gaussian curvature k, then*
>
> $$kA = 2\pi\chi$$

Exercise 12.18 A society of Flatlanders lives on a surface of area 1,984,707 square meters and constant Gaussian curvature -3.1658×10^{-6} meters^{-2}. What is the global topology of their surface? □

Exercise 12.19 Which is more curved, a $P^2 \# P^2 \# P^2$ with area 6 square meters, or a $T^2 \# T^2$ with area 9 square meters? □

You might wonder how Flatlanders measure curvature. They can do it by measuring the area and angles of a triangle. Recall that $A_\triangle = (\alpha + \beta + \gamma) - \pi$ for a triangle on a sphere of radius one, $A_\triangle = \pi - (\alpha$

$+ \beta + \gamma)$ for a triangle on the standard hyperbolic plane, and $\alpha + \beta + \gamma = \pi$ for a triangle on a flat surface. These three formulas can be summarized as $kA_\triangle = (\alpha + \beta + \gamma) - \pi$, where $k = -1$, 0, or +1 according to the geometry. Not surprisingly, this formula is valid for triangles on a surface of any constant Gaussian curvature k. In fact for $k > 0$ the formula is equivalent to the formula $A_\triangle = r^2[(\alpha + \beta + \gamma) - \pi]$ you derived in Exercise 9.3. (By the way, throughout this chapter $A_\triangle$ denotes the area of a triangle while A denotes the area of a whole surface.)

Exercise 12.20 A Flatlander survey team has measured the angles of a triangle as 34.3017°, 62.5633°, and 83.1186°, and they have measured its area as 2.81 km^2. Assuming their universe is homogeneous, what is its Gaussian curvature? (Don't forget to convert the angles to radians.)

The Flatlanders later discover that their universe is orientable and has an area of roughly 250,000 km^2. Deduce the global topology. □

The Gauss–Bonnet formula can be generalized to apply to *non*-homogeneous surfaces whose Gaussian curvature varies irregularly from point to point. The idea is that positive and negative curvature cancel, and the net total curvature equals $2\pi\chi$. On a doughnut surface, for example, the positive curvature cancels the negative curvature exactly. (The outer, convex portion of a doughnut surface has positive curvature, while the portion around the hole is negatively

curved.) No matter how you deform a surface, you can never change its total curvature! When you create positive curvature in one place, you invariably create an equal amount of negative curvature someplace else.

Exercise 12.21 When you make a blip on a surface you create a small region of positive curvature (Figure 12.5). Where is the compensating region of negative curvature? □

One can use the language of calculus to state this latest Gauss–Bonnet formula precisely. The total curvature is represented as the integral of the curvature

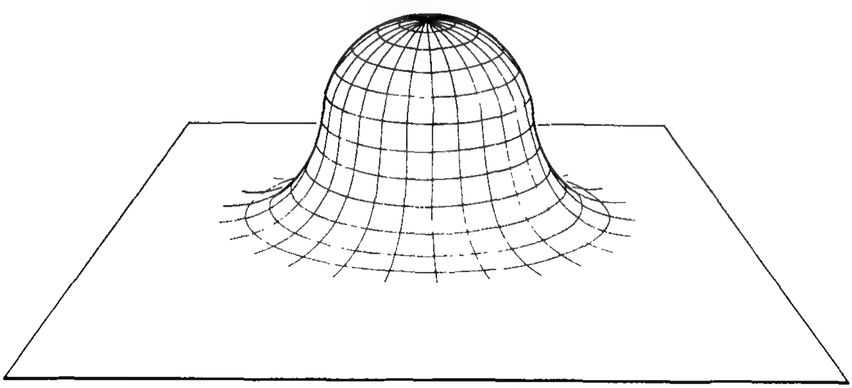


Figure 12.5 When you make a blip in a surface you create a small region of positive curvature. Where is the compensating negative curvature?

over the surface, and the Gauss–Bonnet formula reads

$$\int k\, dA = 2\pi\chi$$

[Mathematically experienced readers can easily justify the steps in the following proof. The basic idea is to (1) use a sufficiently fine cell-division to approximate the integral as a sum, (2) generalize the formula $kA_{\triangle} = (\alpha + \beta + \gamma) - \pi$ to apply to n-gons (see Figure 12.2) and substitute it in, and (3) proceed *exactly* as in the derivation of the elliptic Gauss–Bonnet formula.

$$\begin{aligned}
\int k\, dA &= \sum k_i A_i \\
&= \sum [(\alpha_i + \beta_i + \cdots + \xi_i) - (n_i - 2)\pi] \\
&= \sum (\text{all angles}) - \pi\left(\sum n_i\right) + \pi\left(\sum 2\right) \\
&= 2\pi v - 2\pi e + 2\pi f \\
&= 2\pi\chi
\end{aligned}$$

Note that when k is constant this formula reduces to $kA = 2\pi\chi$.]

Part III

Geometries on Three-Manifolds

This Page Intentionally Left Blank

13

Four-Dimensional Space

In Edwin Abbott's book *Flatland*, A Square's two-dimensional world happens to be embedded in a three-dimensional space. The climax of the book occurs when a sphere from this three-dimensional space comes to visit A Square and tell him about the world of three dimensions. Not surprisingly, A Square finds the sphere's explanations completely unenlightening. So "the sphere, having in vain tried words, resorts to deeds" (*Flatland*, p. 77). He goes over to A Square's locked cupboard, picks up a tablet, moves it over a

little ways, and plunks it back down into the plane of Flatland (Figure 13.1).

Naturally enough, A Square is horrified. As far as he can tell the tablet somehow dematerialized, passed through the wall of his locked cupboard in this condition, and then rematerialized on the other side of the room. This was very, very disconcerting.

Not that we can blame poor A Square. Throughout his whole life he's experienced only forward/backward and left/right motions, so it's not surprising that he has trouble dealing with what we Spacelanders would call up/down motions. When the tablet moves away from the cupboard without moving forwards, backwards, to the right, or to the left, it's only natural for him to assume that it stayed put but somehow became ethereal enough to subsequently pass through the wall (Figure 13.2).

Imagine how you would feel if some strange creature, after rambling on for a while about four-dimensional space, were to remove a pitcher of juice from your fridge and place it on your table—without opening the refrigerator door! The explanation, of course, is that the creature lifts the pitcher "*up*" into the fourth dimension passes it "*over*" the refrigerator wall, and lowers it back "*down*" into our three-dimensional space. Compare Figure 13.3 to Figure 13.2.

In the next chapter we'll use four-dimensional space to define the hypersphere. The hypersphere requires four dimensions for its definition just as an ordinary sphere requires three dimensions. First,

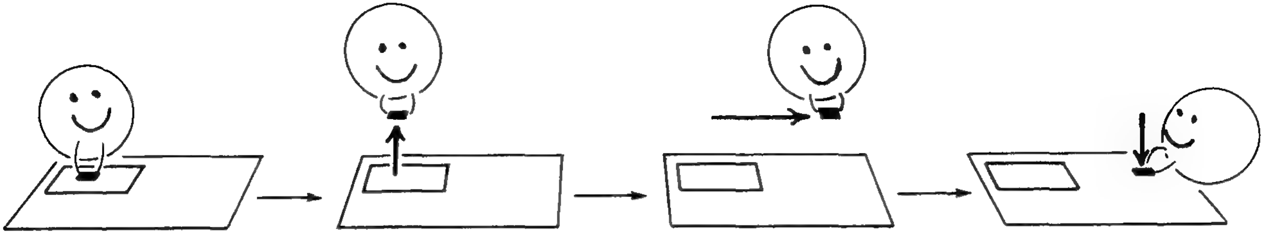


Figure 13.1 How the sphere removed the tablet from the cupboard.

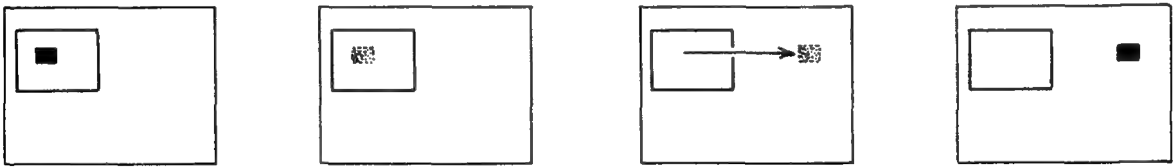


Figure 13.2 A Square's erroneous impression of the incident.

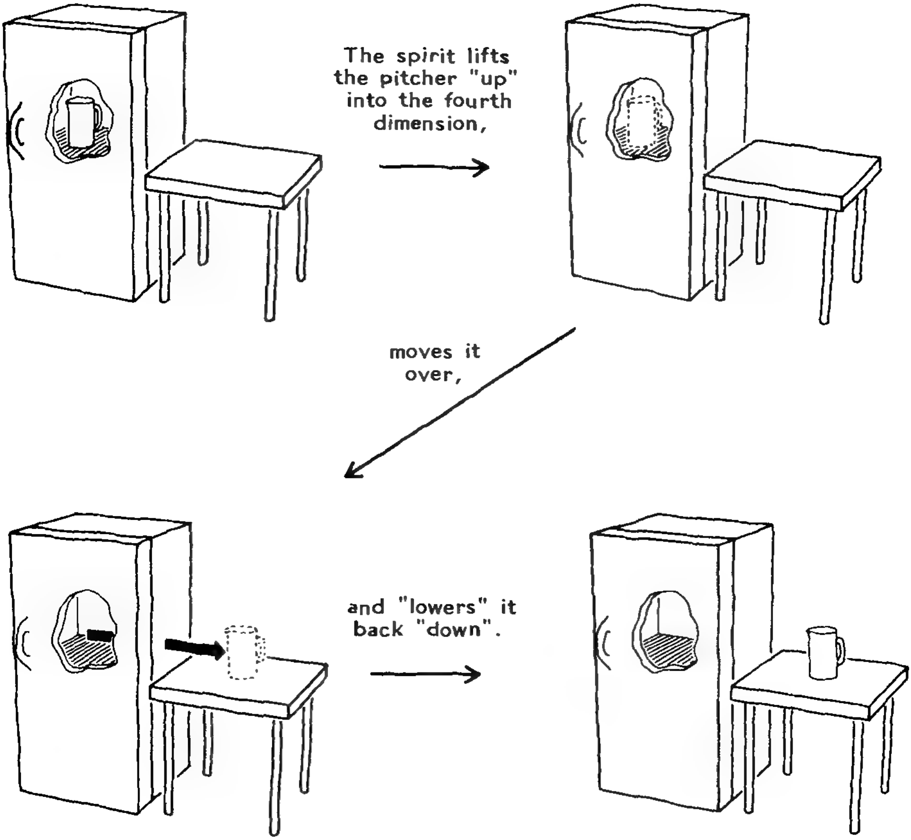


Figure 13.3 How the four-dimensional creature removes the juice pitcher from the refrigerator without opening the door.

though, we'll take a look at some other four-dimensional phenomena, and also deal with a few philosophical questions.

OTHER FOUR-DIMENSIONAL PHENOMENA

On p. 11 of *Geometry, Relativity and the Fourth Dimension*, R. Rucker tells of a mystic name Zöllner who

thought that spirits were four-dimensional beings capable of snatching up three-dimensional objects in much the same way that the sphere snatched up A Square's tablet:

> . . . Prof. Zöllner was also concerned with getting the spirits to do something that would provide a lasting and incontrovertible proof of their four-dimensionality. His idea was a good one. He had two rings carved out of solid wood, so that a microscopic examination would confirm that they had never been cut open. The idea was that spirits, being free to move in the fourth dimension, could link the two rings without breaking or cutting either one. In order to ensure that the rings had not been carved out in a linked position, they were made of different kinds of wood, one alder, one oak. Zöllner took them to a seance and asked the spirits to link them, but unfortunately, they didn't.

Figure 13.4 shows how the spirits were supposed to link the two rings.

Exercise 13.1 How could four-dimensional spirits untie a knotted loop of rope such as the one in Figure 13.5? □

Exercise 13.2 In Chapter 4 we could embed a Klein bottle in E^3 only by allowing it to intersect itself (recall Figure 4.9). Explain how to embed a Klein bottle in E^4 with no self-intersection. ("Ordinary" four-dimensional space is abbreviated E^4, just as ordinary three-dimensional space is E^3, a plane is E^2, and a line is E^1.) □

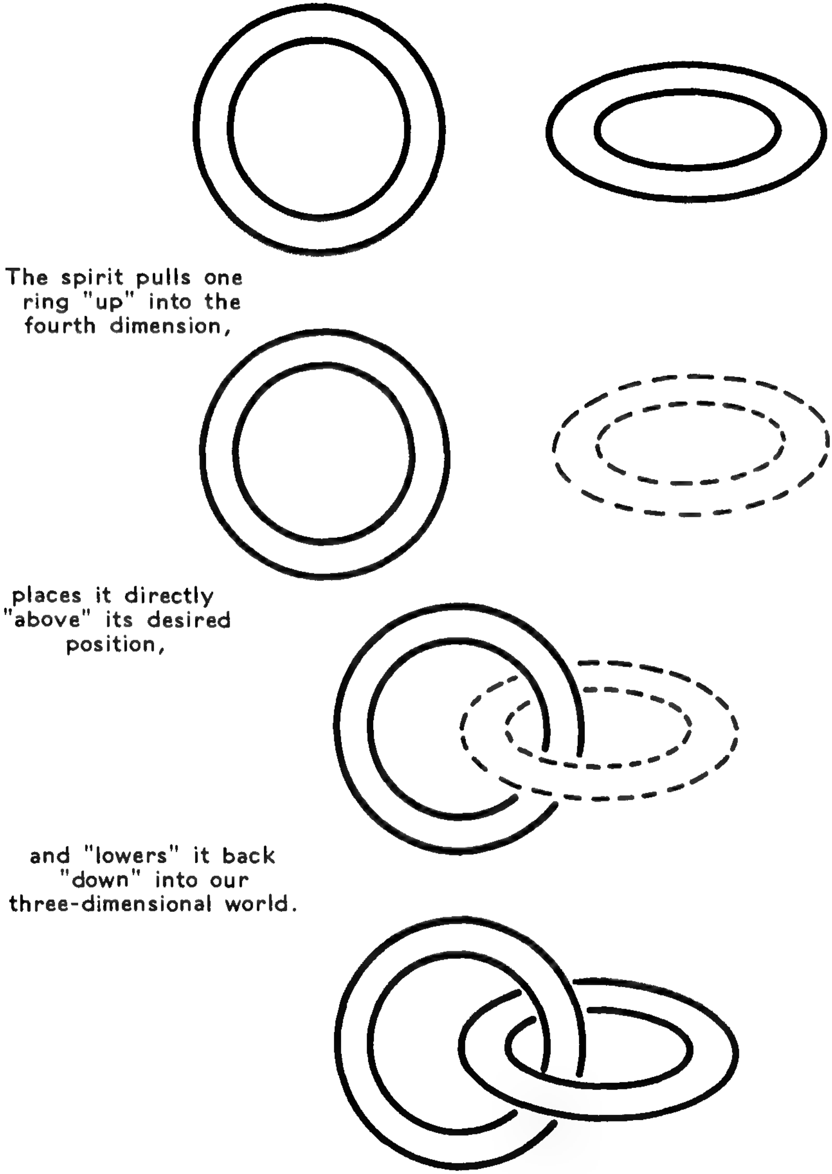


Figure 13.4 A four-dimensional spirit could link two rings together.

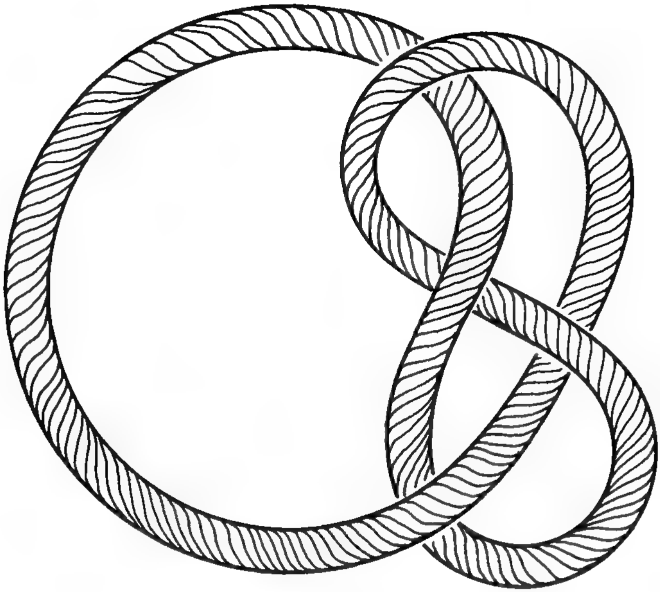


Figure 13.5 How could a four-dimensional spirit untie this "figure-eight" knot?

Before moving on I would like to mention one more four-dimensional curiosity: a sphere can be knotted in four-dimensional space just like a circle can be knotted in three-dimensional space. Give some thought to how A Square imagines a knotted circle (Figure 13.6) and then try to understand the knotted sphere by analogy (Figure 13.7).

PHILOSOPHICAL COMMENTS

(1) Many authors claim that you cannot visualize four-dimensional space. This simply isn't true. (It is true that you must visualize E^4 differently than E^3.) My

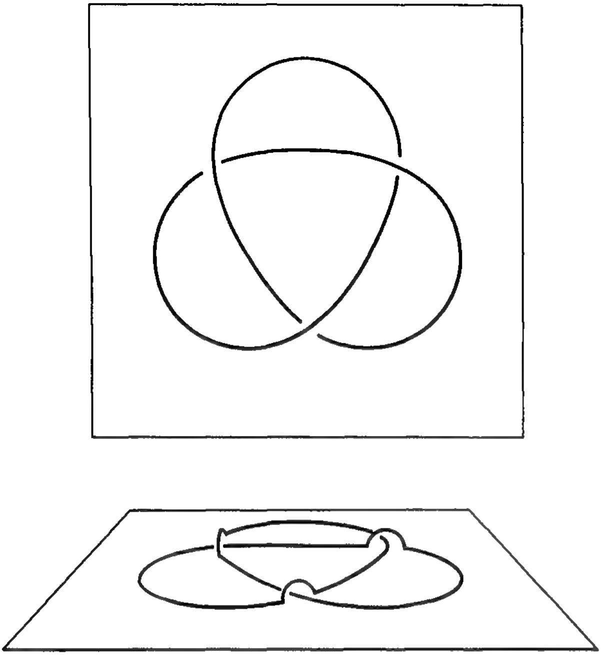


Figure 13.6 To imagine a knotted circle in E^3, A Square first draws a (topological) circle that crosses itself. At each crossing point he then imagines one piece of the circle to pass over the other piece in the third dimension. We Spacelanders can easily visualize the resulting knotted circle.

personal opinion is that your mind is as *capable* of visualizing four dimensions as three. The reason three dimensions is so much easier in practice is that the real universe is three-dimensional: from the day you were born you've been getting practice in understanding three dimensions. At first visualizing four dimensions is difficult and tiring—just as newborn babies no doubt find three dimensions confusing at first. With practice it becomes easier.

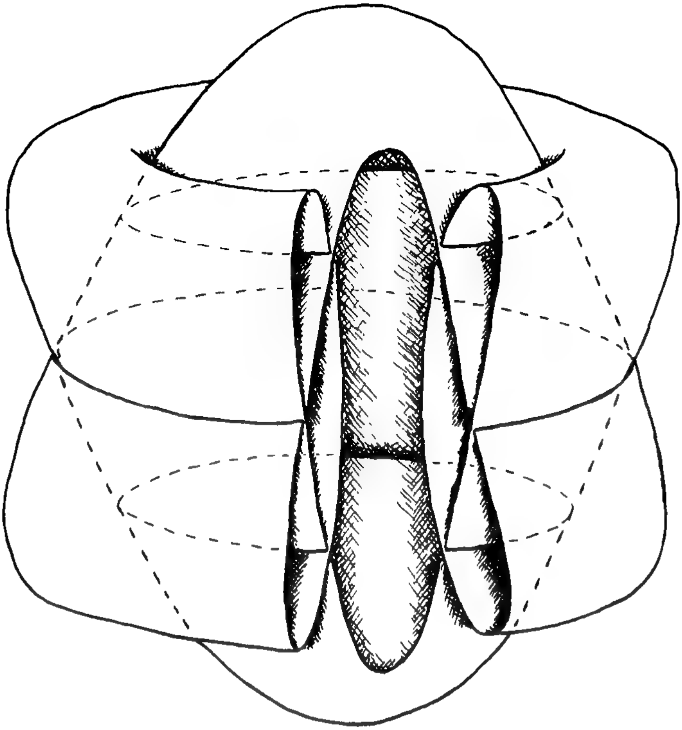


Figure 13.7 To imagine a knotted sphere in E^4, we Spacelanders first draw a (topological) sphere that intersects itself. Along each circle of intersection we then imagine one sheet to pass over the other in the fourth dimension.

(2) Physicists often combine three-dimensional space and one-dimensional time into a four-dimensional entity called *spacetime*. I originally intended to describe this idea in detail here, but I don't think I can improve on the discussion given in Chapter 4, "Time as a Higher Dimension," of R. Rucker's *Geometry, Relativity and the Fourth Dimension*. To give you

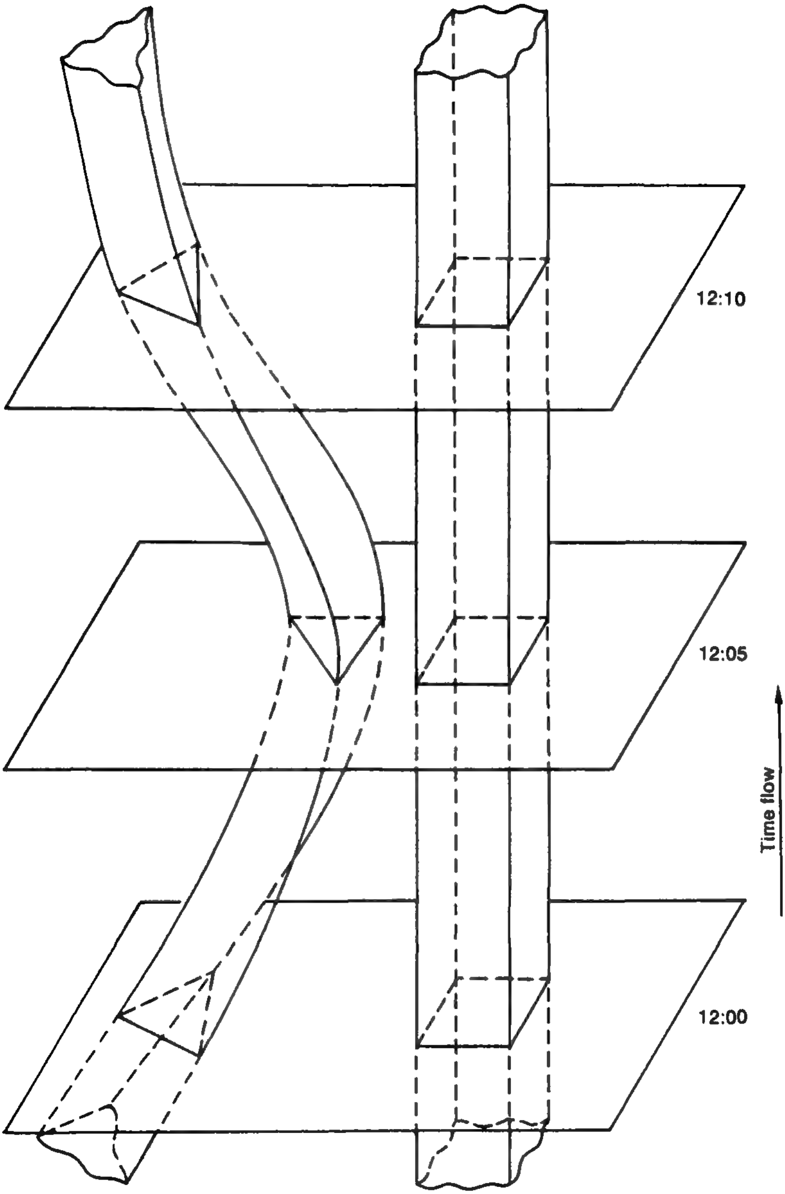


Figure 13.8 Figure 78 from R. Rucker's *Geometry, Relativity and the Fourth Dimension.*

the general idea right here and now, I've reproduced Figure 78 from that chapter (it appears here as Figure 13.8), along with the following explanation:

> To get a good mental image of space-time, let us return to Flatland. Suppose that A. Square is sitting alone in a field. At noon he sees his father, A. Triangle, approaching from the west. A. Triangle reaches A. Square's side at 12:05, talks to him briefly, and then slides back to where he came from. Now, if we think of time as being a direction perpendicular to space, then we can represent the Flatlanders' time as a direction perpendicular to the plane of Flatland. Assuming that "later in time" and "higher in the third dimension" are the same thing, we can represent a motionless Flatlander by a vertical worm or trail and a moving Flatlander by a curving worm or trail, as we have done in Figure 78.
>
> We can think of these 3-D space-time worms as existing timelessly.

Rucker goes on to explain the implications this idea has for consciousness, the perception of time, and the nature of reality. This is thought-provoking stuff—I recommend it highly. (If you're in a hurry, you can easily read that one chapter independently of the rest of the book.)

In Chapter 14 we'll stick to imagining four dimensions in a purely spatial way—time won't enter the picture at all.

(3) A final metaphysical comment: The four-dimensional space we imagine is purely a mathematical abstraction. I make no claim that it exists physically like the three-dimensional universe does.

This Page Intentionally Left Blank

14

The Hypersphere

Roughly speaking, a sphere is like a circle, only one dimension bigger. A hypersphere, or three-sphere, is the analogous three-manifold one dimension bigger than a sphere. Compare the formal definitions of the circle, the sphere, and the hypersphere:

1. The unit circle (= one-sphere = S^1) is the set of points in E^2 that are one unit away from the origin. Anything topologically equivalent to it is called a *topological circle* and anything

geometrically equivalent is called a *geometrical circle.*

2. The unit sphere (= two-sphere = S^2) is the set of points in E^3 that are one unit away from the origin. Anything topologically equivalent to it is called a *topological two-sphere* and anything geometrically equivalent is called a *geometrical two-sphere.*
3. The unit hypersphere (= three-sphere = S^3) is the set of points in E^4 that are one unit away from the origin. Anything topologically equivalent to it is called a topological three-sphere and anything geometrically equivalent is called a geometrical three-sphere.

A Warning on Terminology: Our two-sphere is defined in three-dimensional space, where it is the boundary of a three-dimensional ball. This terminology is standard among mathematicians, but not among physicists. So don't be surprised if you find people calling the two-sphere a three-sphere. They're interested in the dimension of the space that the two-sphere happens to be in, while we're interested in the intrinsic dimension of the two-sphere itself. (A two-sphere is still intrinsically two-dimensional even if it's sitting in E^4, like the knotted two-sphere of Figure 13.7.) Similar comments apply to the three-sphere. Also note carefully the distinction between "sphere" and "ball" as used above. Some people use "sphere" to mean "ball," not us. We do, however, use "disk" and

"ball" interchangeably ("disk" sounds better in two dimensions and "ball" sounds better in three, but the concept is essentially the same).

How can one visualize a three-sphere? To answer this question, let's take a look at how A Square might visualize a two-sphere. Assume for the moment that he's interested only in its topology and not its geometry. He can make things easy for himself by flattening the sphere into the plane of Flatland, as shown in Figure 14.1. The northern and southern hemispheres each become a disk in Flatland. The two disks are superimposed and joined together along their circular boundary (the equator). Thus, the task of visualizing the two-sphere has been reduced to the task of visualizing two superimposed disks. A Square has to remember, though, that the "crease" at the equator doesn't exist in the real two-sphere—it's merely an artifact of the flattening process. This method of visualizing the two-sphere is called the *double disk method*, as is the analogous method of visualizing the three-sphere (see Exercise 14.1).

Exercise 14.1 Imagine a three-sphere topologically as two superimposed solid balls in E^3. These balls are joined together along their spherical boundary. (The equator of S^3 is a two-sphere!) □

Exercise 14.2 To recover the geometry of the two-sphere, A Square imagines one disk bending upward into the third dimension and the other disk bending

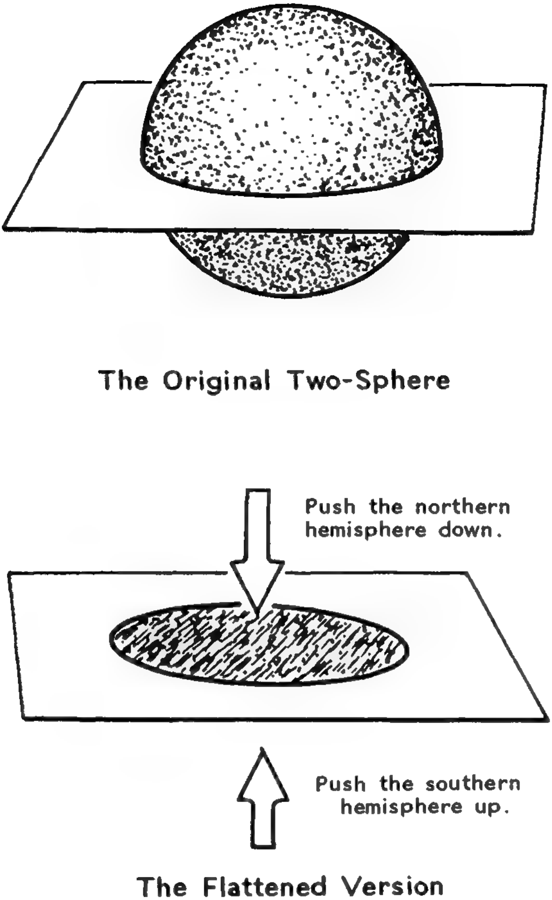


Figure 14.1 A Square visualizes the two-sphere topologically as two superimposed disks.

downward. Modify your mental image from Exercise 14.1 to let one solid ball bend "upward" into the fourth dimension, and the other ball bend "downward." □

Note: Some readers may understand the three-sphere's global topology more easily by imagining it as

two *non*superimposed solid balls whose surfaces are abstractly glued together (Figure 14.2).

Exercise 14.3 Pretend you live in a fairly small three-sphere. What will eventually happen to you if you keep blowing air into an easily stretchable balloon? (Hint: What will eventually happen to A Square in Figure 14.3?) □

Figure 14.4 shows three great circles on a two-sphere (they happen to be the intersection of the two-sphere with each of the coordinate planes). We Spacelanders visualize these great circles quite easily because we can draw the two-sphere in three-dimensional space. Flatlanders, on the other hand, draw the two-sphere flattened into their plane. The equator is still a circle, but the meridians look like line segments. The Flatlanders must remember that each apparent line segment is really two line segments, one arching upward into the third dimension and the other arching downward. Two line segments together form a perfect geometrical circle.

Figure 14.5 shows four "great two-spheres" on a three-sphere (they happen to be the intersection of the three-sphere with each of the coordinate hyperplanes). It's hard for us Spacelanders to visualize the great two-spheres because the three-sphere they lie in has been flattened down into our three-dimensional space. Only the "equatorial" two-sphere still looks like a sphere. Each of the others now looks like a disk. Each is really two disks, of course, one arching up-

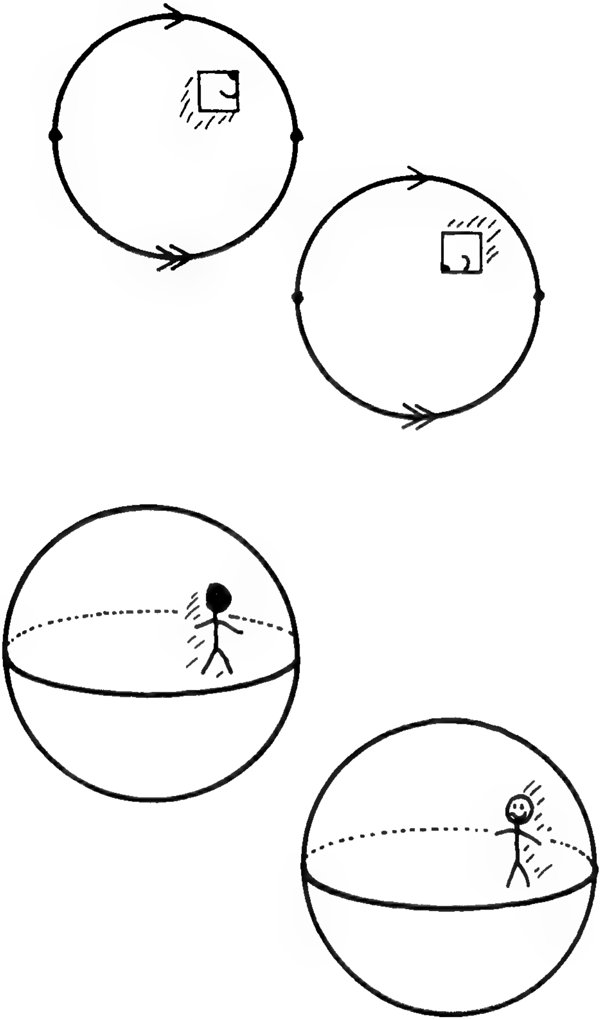


Figure 14.2 A two-sphere may be represented topologically as two disks with edges glued together. Similarly, a three-sphere may be represented topologically as two solid balls with surfaces glued together.

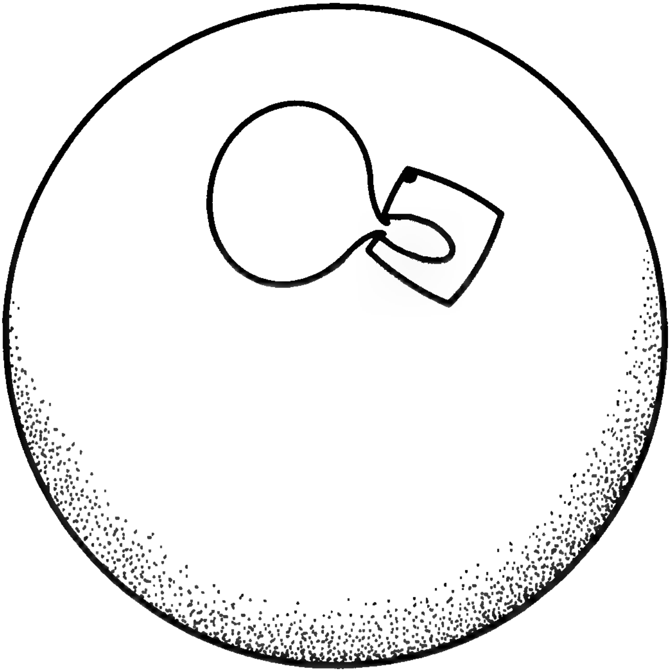


Figure 14.3 A Square inflates a balloon on a two-sphere.

ward into the fourth dimension and the other arching downward. Two disks together form a perfect geometrical sphere.

The preceding two paragraphs show how every slice of a three-sphere is a two-sphere, just as every slice of a two-sphere is a circle.

If we do not allow it to stretch, a piece of a three-sphere will split open in Euclidean space (Figure 14.6) just as a piece of a two-sphere splits open in the plane (Figure 9.8).

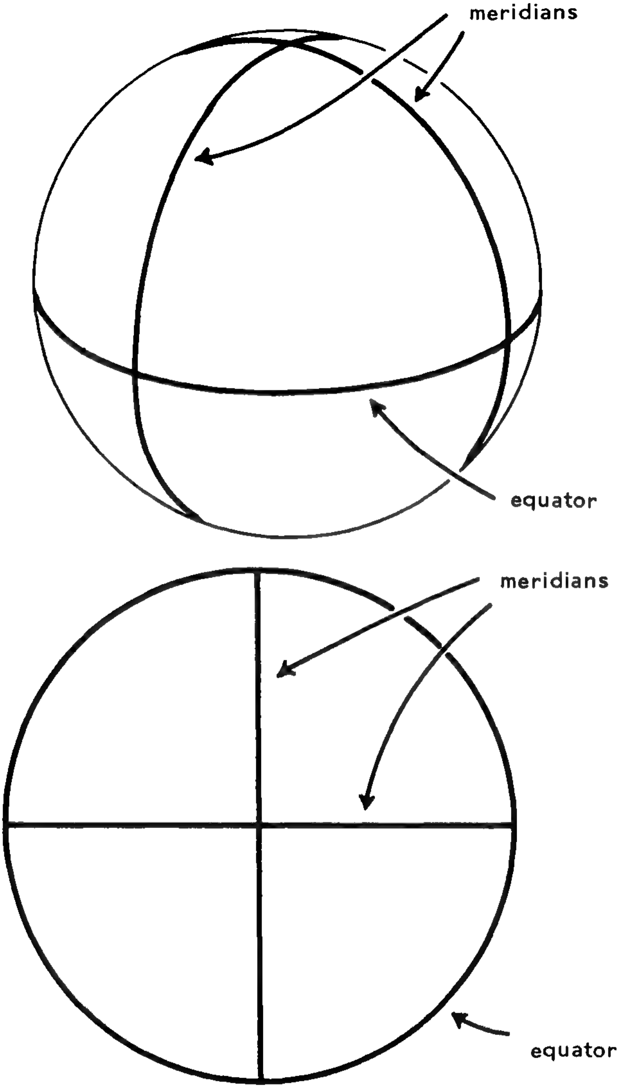


Figure 14.4 The first drawing is a Spacelander's view of three great circles on a two-sphere. The second drawing is a top view of the two-sphere after it has been flattened into a horizontal plane. Only the equator is still a circle. Each meridian has been flattened into a line segment.

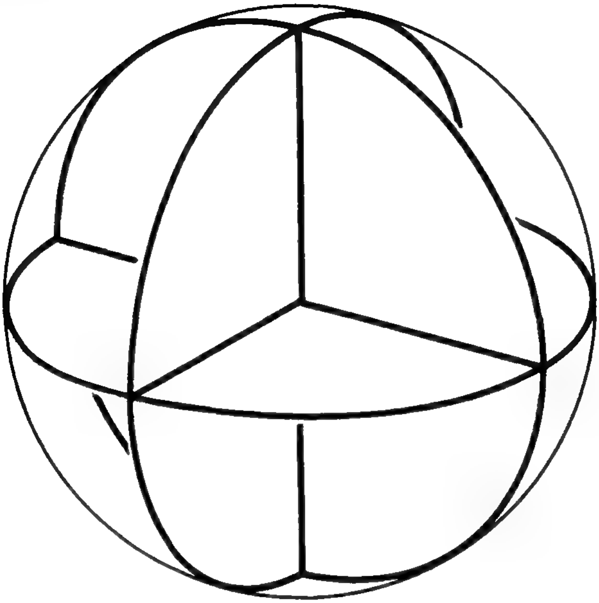


Figure 14.5 This drawing represents four "great two-spheres" in a hypersphere. It is completely analogous to the second drawing in Figure 14.4. The surface of the ball is the three-sphere's "equator." Each disk is the flattened remains of a two-sphere, so imagine each disk as two disks, one bending "upward" and the other "downward" into the fourth dimension.

Polyhedra in a three-sphere have larger angles than do polyhedra in Euclidean space (Figure 14.7). It turns out that polyhedra in "hyperbolic space" (Chapter 15) have smaller angles than do polyhedra in Euclidean space. We'll utilize these facts to find homogeneous geometries for certain three-manifolds (Chapter 16), just as we found homogeneous geometries for surfaces (Chapter 11).

Figure 14.6 If we do not allow it to stretch, a piece of a three-sphere will split open in Euclidean space. Note that every cross-section of this split open ball is a split open disk like the one in Figure 9.8.

Say our universe is a three-sphere. We can measure its curvature by measuring the curvature of a great two-sphere. And we can measure the curvature of a great two-sphere by measuring the angles and area of a triangle lying on it. If we don't care which great two-sphere we are measuring—and we don't because they are all the same—then we can measure any triangle we want.

Exercise 14.4 Gauss measured the angles of a triangle roughly 100 km on a side, probably using in-

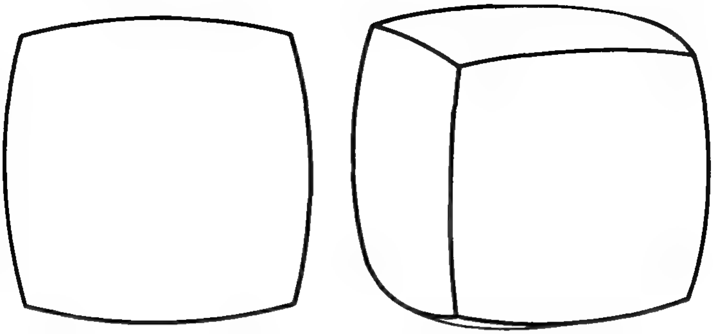


Figure 14.7 Polyhedra in a three-sphere have larger angles than do polyhedra in ordinary Euclidean space, just as polygons on a two-sphere have larger angles than do polygons in the Euclidean plane. Note: When representing a spherical polygon in the Euclidean plane it is customary to make the sides bulge so that the angles come out right. On the sphere itself, of course, the polygon's sides do not bulge—they are perfect geodesics. Similarly, when representing a polyhedron from the three-sphere it is customary to make the sides bulge so that the edge and corner angles come out right, even though the sides do not bulge in the three-sphere itself.

struments accurate to, say, 10 minutes of arc. What is the smallest curvature he could detect? What is the radius of a two-sphere with this curvature? The radius of any great two-sphere is the same as the radius of the three-sphere it lies in, so the answer to the previous question represents the radius of the largest S^3-universe whose curvature Gauss could detect. Note that small three-spheres have large curvature and large three-spheres have small curvature, so it's easier to detect the curvature of a small S^3-universe than a large one. Do you think it's likely that the universe is this small?

I'm sure Gauss was well aware of the above considerations. He surely wasn't trying to measure the curvature of the universe, but was instead interested in measuring the curvature of Earth's ellipsoidal (not spherical!) surface. By the way, modern cosmologists attempt to measure the curvature of the universe by very different means (Part IV). □

PROJECTIVE THREE-SPACE

Back in Chapter 4 you constructed the projective plane (P^2) by gluing together opposite points on the circular rim of a hemisphere of S^2 (recall Figure 4.11). The hemisphere's local geometry matched up nicely across the "seam" (Figure 4.12), so you got a surface with the same local geometry as S^2, but a different global topology.

You can make *projective three-space* (P^3) in the same way. Start with a hemisphere of S^3 and glue together opposite points on its (spherical) boundary. It should be clear (at least from analogy) that the hemisphere's local geometry matches up nicely across the spherical "seam." Thus P^3 is a three-manifold with the same local geometry as S^3, but a different global topology.

In Chapter 4 we noted that for topological purposes we can think of P^2 as a disk with opposite boundary points glued, and we can think of P^3 as a

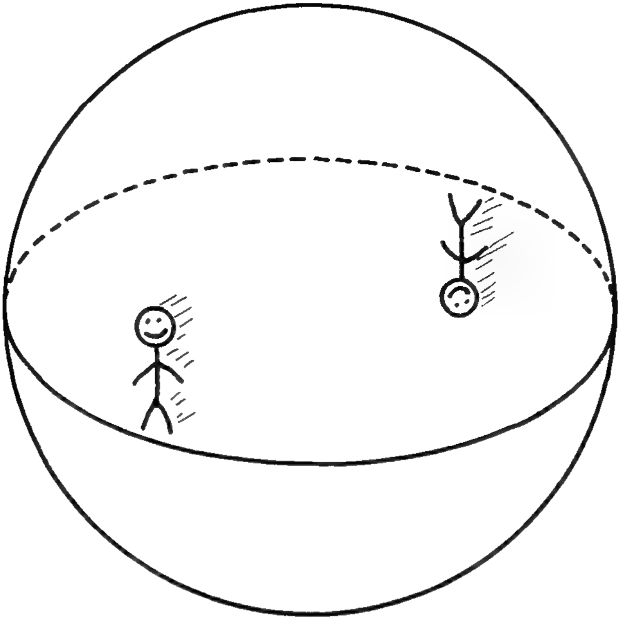


Figure 14.8 When you cross the "seam" of P^3 you come back with your head where your feet were and your left side where your right side was. In effect you rotate a half-turn.

solid ball with opposite boundary points glued. The projective plane is nonorientable: when a Flatlander crosses the "seam" he comes back left-right reversed. Projective three-space, on the other hand, *is* orientable. When you cross the "seam" you come back both left-right reversed *and* top-bottom reversed. In effect, you get mirror-reversed two ways, so you come back as your old self! The only difference is that you've been rotated 180°. (See Figure 14.8.)

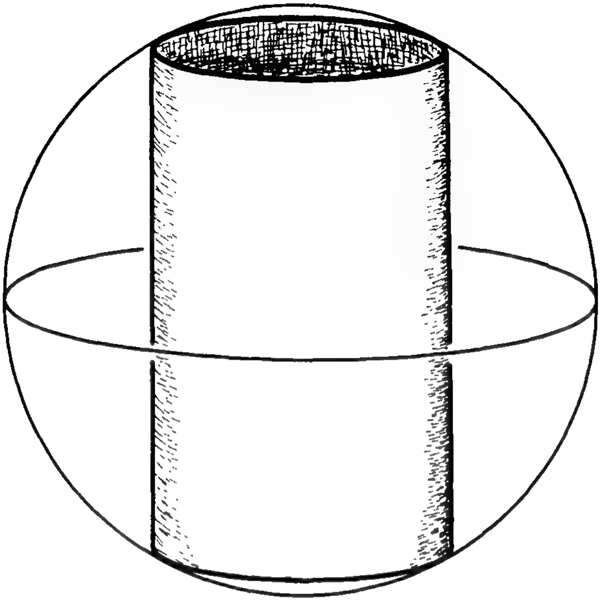


Figure 14.9 Does the cylinder become a torus or a Klein bottle in projective three-space?

Exercise 14.5 See Figure 14.9. Does the cylinder form a torus or a Klein bottle in projective three-space? Is it orientable? Is it two-sided? □

Exercise 14.6 Find a copy of P^2 embedded in P^3. Is it orientable? Is it two-sided? □

15

Hyperbolic Space

Hyperbolic space is just like the hyperbolic plane, only one dimension bigger. In fact, every two-dimensional slice of hyperbolic space is a hyperbolic plane (Figure 15.1) in the same way that every two-dimensional slice of Euclidean space is a Euclidean plane and every two-dimensional slice of a hyperspace is a two-sphere. Hyperbolic space is homogeneous. It is often abbreviated as H^3.

A polyhedron in H^3 has smaller angles than a polyhedron in Euclidean space (Figure 15.2). This fact

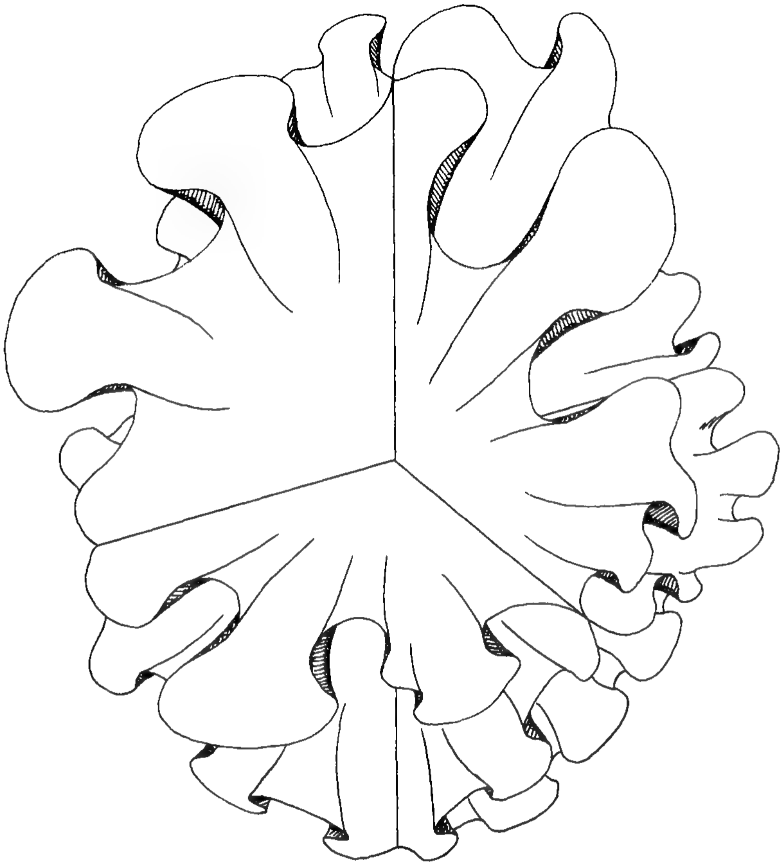


Figure 15.1 Every slice of hyperbolic space is a hyperbolic plane.

will be crucial in the next chapter when we find homogeneous geometries for certain three-manifolds.

Figure 15.3 shows a series of successively larger hyperbolic triangles. Note that larger triangles have smaller angles. The figure sheds some light on why no

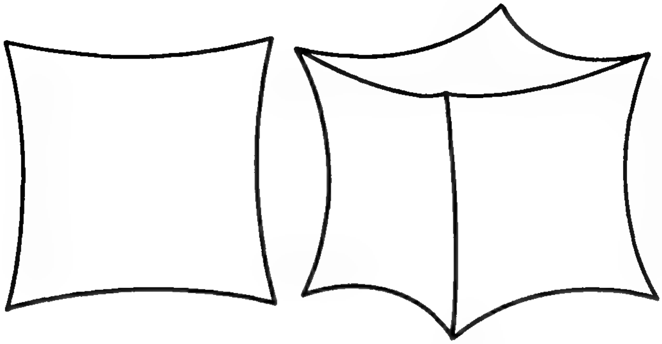


Figure 15.2 Polyhedra in hyperbolic space have smaller angles than do polyhedra in Euclidean space, just as polygons in the hyperbolic plane have smaller angles than do polygons in the Euclidean plane. Note: When representing a "hyperbolic polygon" in the Euclidean plane it is customary to make the sides bend inward so that the angles come out right. In the hyperbolic plane itself, of course, the polygon's sides do not bend inward—they are perfect geodesics. Similarly, when representing a polyhedron from hyperbolic space it is customary to make the sides bend inward so that the edge and corner angles come out right, even though the sides do not bend inward in hyperbolic space itself.

triangle in the (standard) hyperbolic plane can have area greater than π. When you try to draw a triangle with more area you find that the triangle's sides don't meet. In other words, you end up with three nice straight geodesics, no two of which intersect! The hyperbolic plane has some strange properties.

Imagine that you have encountered a party of extracosmic aliens. The aliens come from a highly curved hyperbolic universe, and are in our universe only for a visit. They have obviously mastered the secrets of interuniversal travel. You have heard that hy-

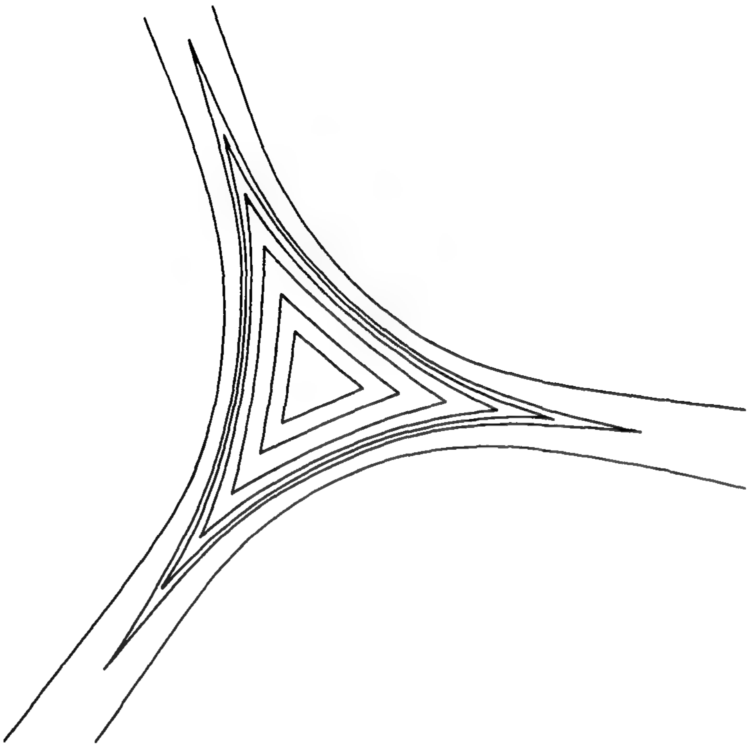


Figure 15.3 A series of successively larger hyperbolic triangles represented in the Euclidean plane. As explained in the caption to Figure 15.2, the triangles' sides don't really bend inward in the hyperbolic plane itself, but the angles shown here are correct.

perbolic universes are somehow more spacious than Euclidean ones, so you ask the aliens to take you home with them. They oblige. When you get to their universe you look out into the sky at a distant galaxy. The light reaching your eyes naturally travels along nice straight geodesics, but in hyperbolic space geo-

desics do weird things. The light reaches your left eye at a slightly different angle than it reaches your right eye: you have to look somewhat cross-eyed to focus on the galaxy! (See Figure 15.4.) Your brain, used to interpreting visual data in an approximately Euclidean universe, decides that since you have to look cross-eyed the galaxy must be very close. In fact everything in the hyperbolic universe seems to be within a few meters of you. Even though it's really very spacious, a hyperbolic universe can *appear* very cramped.

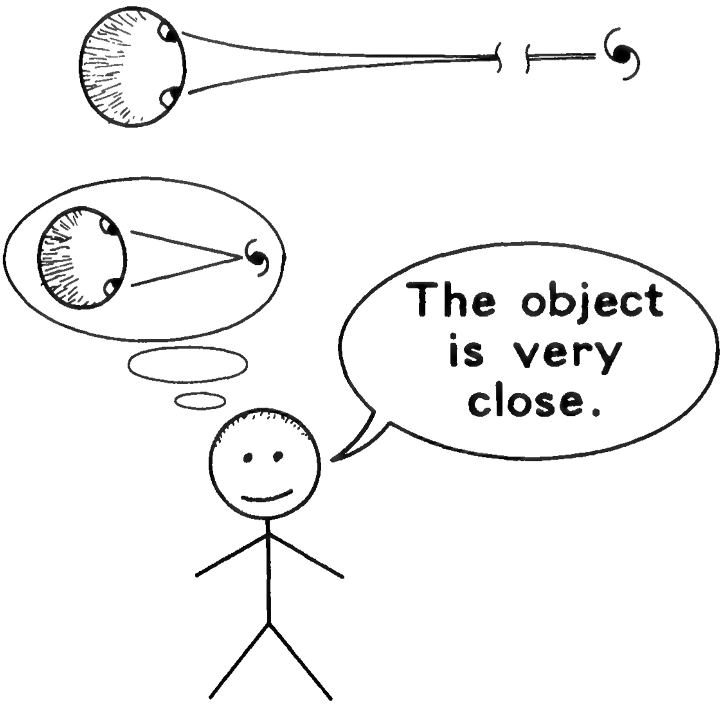


Figure 15.4 According to your binocular vision, very distant objects look close.

Exercise 15.1 Since the effects described in the preceding paragraphs have not been observed in the real universe, does this mean that it cannot have the geometry of H^3? □

Exercise 15.2 How would an H^3-universe appear to its own inhabitants? (Hint: It needn't appear the same to them as it would to an outsider!) □

If you have a pair of red-blue glasses, you can explore hyperbolic space in stereoscopic 3-D using interactive 3-D graphics software available for free at *www.northnet.org/weeks/SoS*.

16

Geometries on Three-Manifolds I

The *Seifert–Weber space* consists of a dodecahedron whose opposite faces are glued with three-tenths turns (Figure 16.1). This three-manifold fails to have a Euclidean geometry for essentially the same reason that the third surface in Figure 11.1 failed to have one: the dodecahedron's twenty corners all come together at a single point, and they are much too fat to fit together properly. The solution is the same as in Chapter 11. Put the dodecahedron in hyperbolic space and let it expand until its corners are skinny enough that they do fit together (Figure 16.2). A Seifert–We-

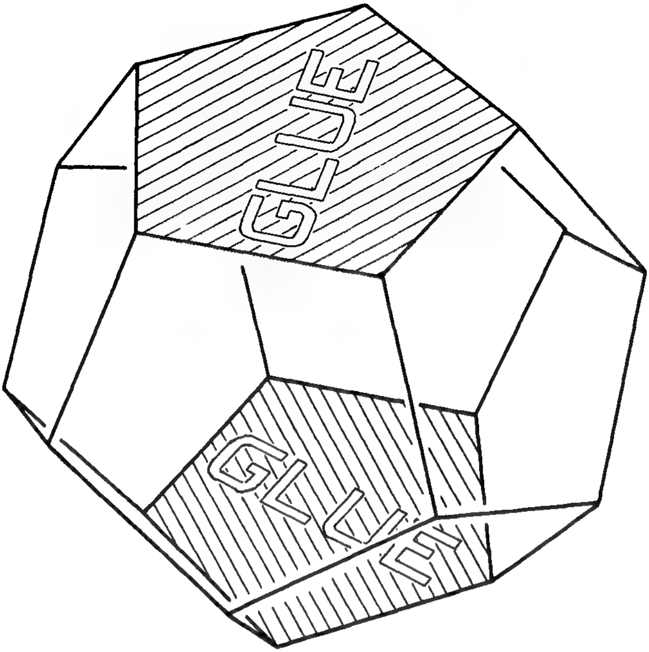


Figure 16.1 In the Seifert–Weber space every face of the dodecahedron is glued to the opposite face with a three-tenths clockwise turn. [A technical point: You might think that gluing the top to the bottom with a clockwise turn would be the same as gluing the bottom to the top with a counterclockwise turn, but this is not the case. Study the figure and you will see that gluing the top to the bottom with a clockwise turn (as viewed from above) works out the same as gluing the bottom to the top with a clockwise turn (as viewed from below). Thus the description of the Seifert–Weber space is self-consistent.]

ber space made from the appropriate dodecahedron has a homogeneous hyperbolic geometry.

The *Poincaré dodecahedral space* consists of a dodecahedron whose opposite faces are glued with one-tenth turns (Figure 16.3). This three-manifold fails to

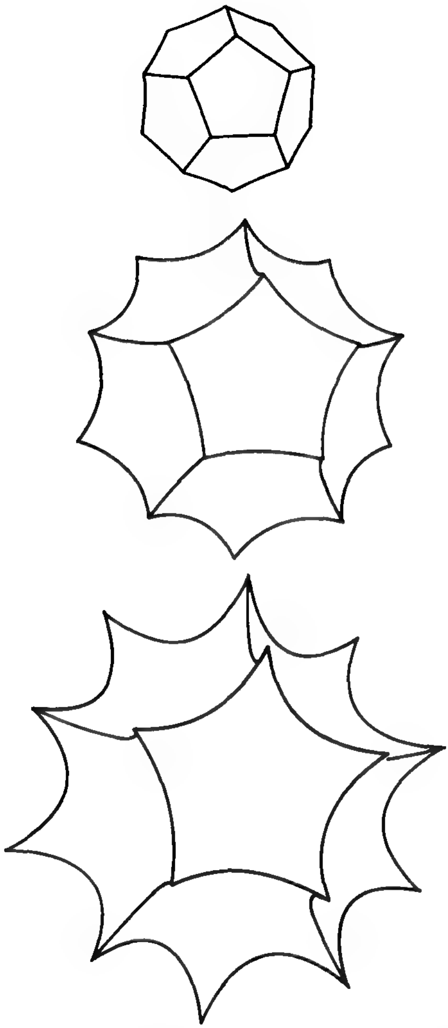


Figure 16.2 Let a dodecahedron expand in hyperbolic space until its corners are the right size to all fit together at a single point. The angles shown here are accurate, but in hyperbolic space itself the dodecahedron's faces do not bend inward.

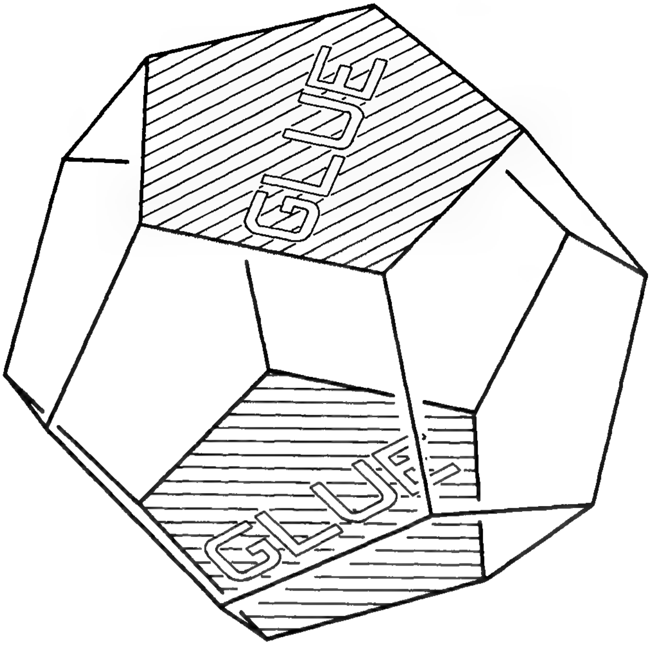


Figure 16.3 In the Poincaré dodecahedral space every face of a dodecahedron is glued to the opposite face with a one-tenth clockwise turn.

have a Euclidean geometry for essentially the same reason that the *first* surface in Figure 11.1 failed to have one: the dodecahedron's twenty corners come together in five groups of four corners each, and they are a little too skinny to fit together properly. The solution is the same as in Chapter 11. Put the dodecahedron in a hypersphere and let it expand until its corners are fat enough that they do fit together (Figure 16.4). A Poincaré dodecahedral space made from the appropriate dodecahedron has the homogeneous

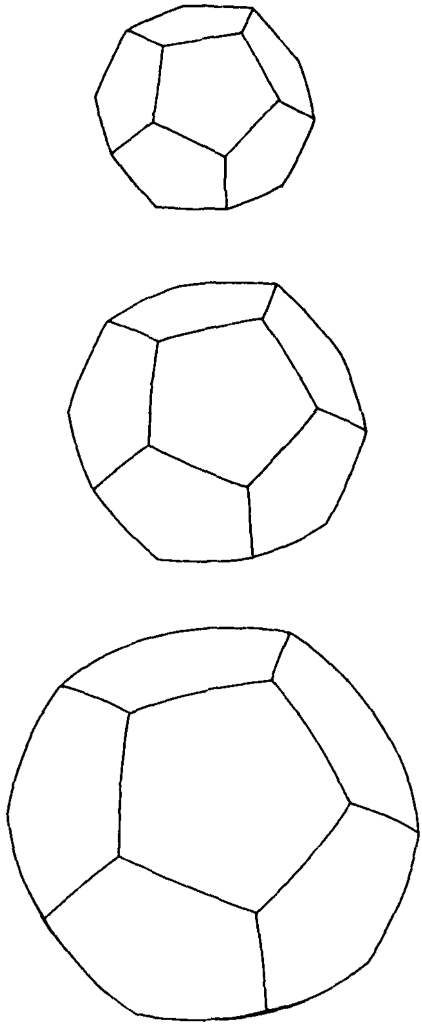


Figure 16.4 Let a dodecahedron expand in a three-sphere until its corners are the right size to fit together in groups of four. The angles shown here are accurate, but in the three-sphere itself the dodecahedron's faces do not bulge outward.

geometry of the hypersphere (i.e. three-dimensional elliptic geometry).

Both the Seifert–Weber space and the Poincaré dodecahedral space appeared in 1933 in C. Weber and H. Seifert's article "The two dodecahedral spaces" (Die beiden Dodekaederräume, Mathematische Zeitschrift, Vol. 37, no. 2, p. 237). The Poincaré dodecahedral space is named in honor of Henri Poincaré (pronounced "pwan-ka-RAY") because it is topologically the same as a three-manifold Poincaré discovered in the 1890s. Poincaré, though, didn't know that his manifold could be made from a dodecahedron! He was interested in it because it had certain properties in common with the hypersphere, namely the same "homology." (He had previously thought that the only three-manifold with the homology of the three-sphere was the three-sphere itself.)

At this point it's appropriate to note that in the three-torus a cube's eight corners all come together at a single point, and they fit perfectly (Figure 16.5). This is why the three-torus has Euclidean geometry. By the way, all the other three-manifolds in Chapter 7 have Euclidean geometry for the same reason.

We have seen that the Seifert–Weber space admits three-dimensional hyperbolic geometry, the Poincaré dodecahedral space admits three-dimensional elliptic geometry, and the three-torus admits three-dimensional Euclidean geometry. It would be nice if every three-manifold admitted one of these three geometries, but the actual situation is not that simple.

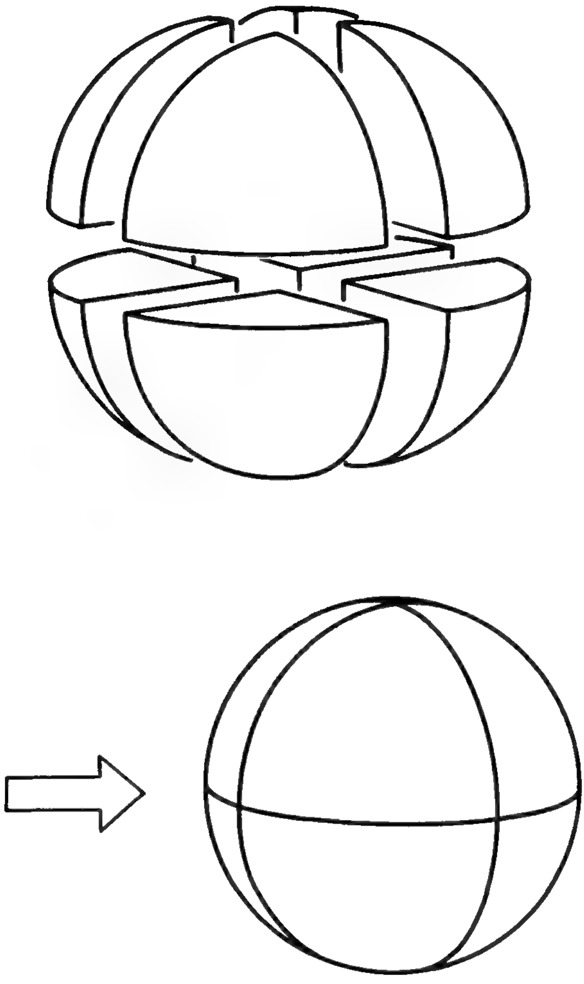


Figure 16.5 Eight corners of a cube fit together perfectly just as they are.

For example, $S^2 \times S^1$ has a homogeneous geometry different from the three just mentioned. It was only in the 1970s that people have come to understand the situation more fully. Chapter 18 explains what is known and/or conjectured. Chapter 17 provides the examples necessary for Chapter 18. Chapters 19–22 discuss the nature of the universe, drawing on what you now know about three-manifolds. Chapters 19–22 do not depend on Chapters 17 or 18, so you can read them immediately if you want.

Exercise 16.1 The *tetrahedral space* is a tetrahedron with faces glued as indicated in Figure 16.6(a). How do the tetrahedron's corners fit together, i.e. how many groups of how many corners each? (Hint: Start in one corner and see which of the other three corners you can reach by passing through a face.) Do the corners have to expand or shrink to fit properly? What homogeneous geometry does this manifold admit? □

Exercise 16.2 The *quaternionic manifold* is a cube with each face glued to the opposite face with a one-quarter clockwise turn (Figure 16.6(b)). How do the cube's corners fit together? What homogeneous geometry does this manifold admit? By the way, the manifold's funny name arises from the fact that its symmetries can be modelled in the quaternions, a number system like the complex numbers but with three imaginary quantities instead of just one. □

You can explore the 3-manifolds of this chapter using interactive 3-D graphics software available for free at *www.northnet.org/weeks/SoS*.

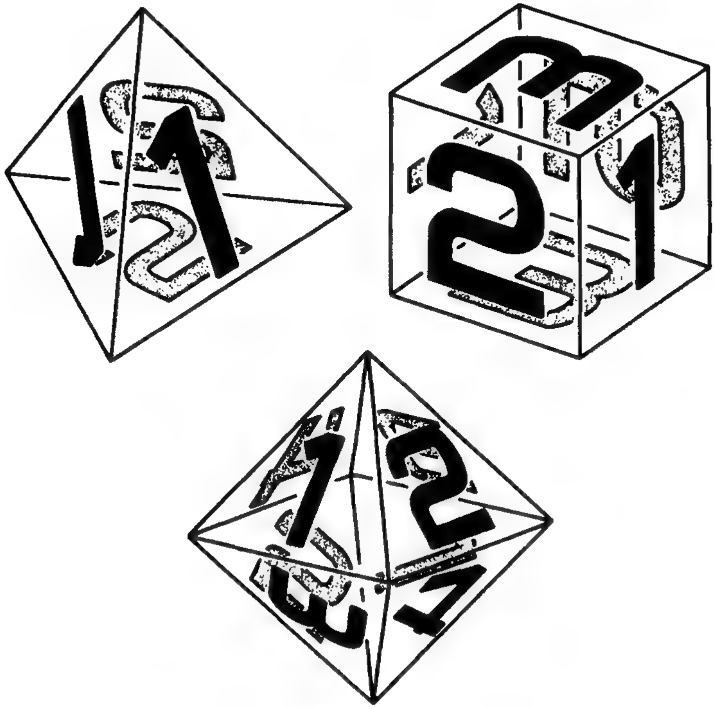


Figure 16.6 The manifolds for Exercises 16.1, 16.2, and 16.3.

Exercise 16.3 The *octahedral space* is an octahedron with each face glued to the opposite face with a one-sixth clockwise turn (Figure 16.6(c)). Find a homogeneous geometry for the octahedral space. (This exercise is a little harder than the preceding two. Even after you figure out how the corners fit, it's still not obvious whether they are too fat, too skinny, or just right. You can work it out by elementary means, but you have to get your hands dirty.) □

17

Bundles

A cylinder is the product of an interval and a circle because it is both an interval of circles and a circle of intervals (Figure 17.1). (For a review of products, see Chapter 6.)

A Möbius band is also a circle of intervals (Figure 17.2), but it fails to be an interval of circles. It is almost a product, but not quite. It therefore qualifies as an interval bundle over a circle. In general a *bundle* over a circle is a bunch of things smoothly arranged in a circle, whether or not they form a product. For example, the quarter turn manifold from Exercise 7.3

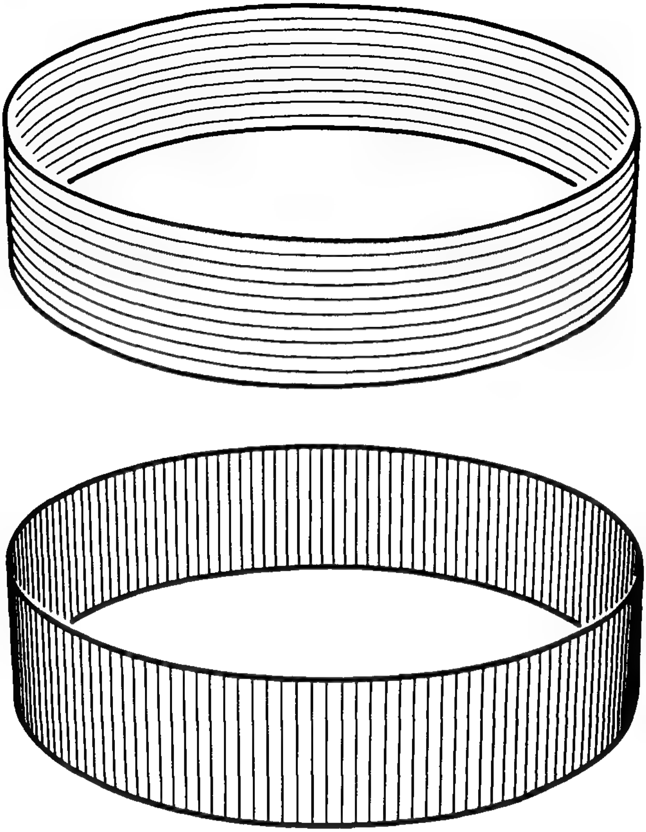


Figure 17.1 A cylinder is both an interval of circles and a circle of intervals.

is a torus bundle over a circle. Figure 17.3 reviews the construction of the quarter turn manifold: the front and back, and left and right, faces of a cube are glued in the straightforward way, but the top is glued to the bottom with a quarter turn. Figure 17.4 will help you understand the manifold's global topology. Every flat three-manifold in Chapter 7 was either a torus bundle

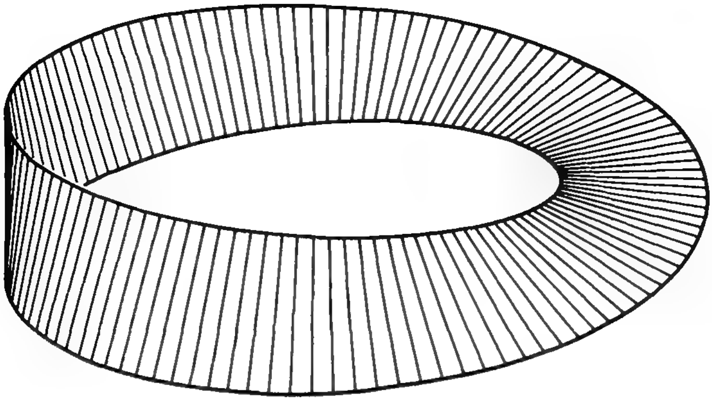


Figure 17.2 A Möbius band is almost the product of an interval and a circle.

over a circle (T^2-bundle over S^1) or a Klein bottle bundle over a circle (K^2-bundle over S^1).

Exercise 17.1 The ⅓ turn manifold, introduced in Exercise 7.13, is a (hexagonal) torus bundle over a circle. Draw a picture of it analogous to the picture of the ¼ turn manifold in Figure 17.4. Do the same for the ⅙ turn manifold. □

The ¼ turn manifold, the ⅓ turn manifold and the ⅙ turn manifold can each be represented as a circle of tori in three-dimensional space, as in Figure 17.4. On the other hand, when the top of a cube or prism is glued to the bottom with a side-to-side flip, then you cannot physically carry out the gluing in three-dimensional space to get a picture like Figure 17.4 and you must fall back to a picture like Figure 17.3 to understand the bundle more abstractly.

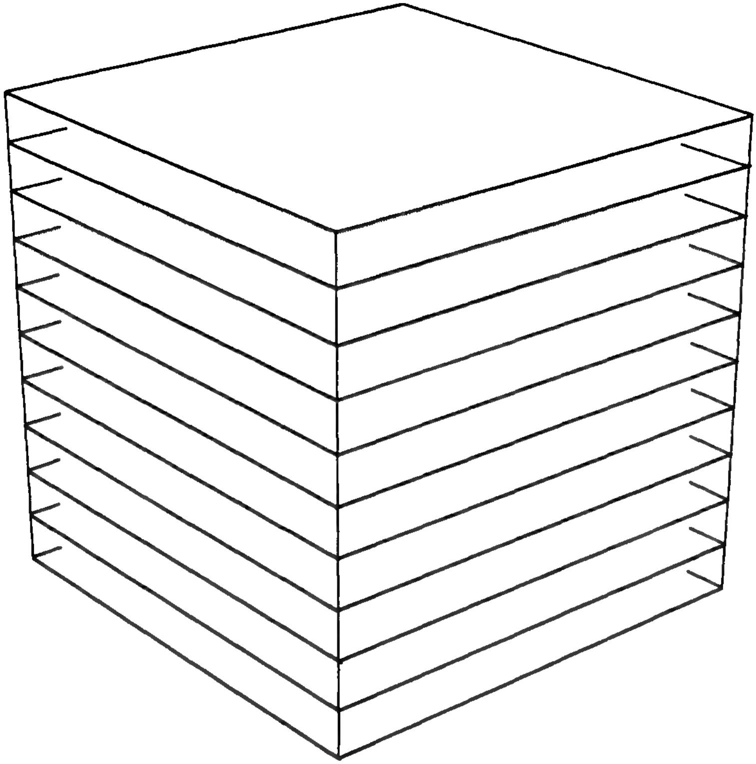


Figure 17.3 Opposite sides of the cube are glued in the straightforward way, but the top is glued to the bottom with a quarter turn. Each horizontal layer forms a torus.

Exercise 17.2 $K^2 \times S^1$ is both a T^2-bundle over S^1 *and* a K^2-bundle over S^1. Draw one picture representing it as a circle of tori, and another representing it as a circle of Klein bottles. Your pictures should be analogous to Figure 17.3 rather than 17.4. (You can draw a picture analogous to Figure 17.4 for the K^2-bundle but not for the T^2-bundle.) Figure 17.5 reviews the construction of $K^2 \times S^1$. □

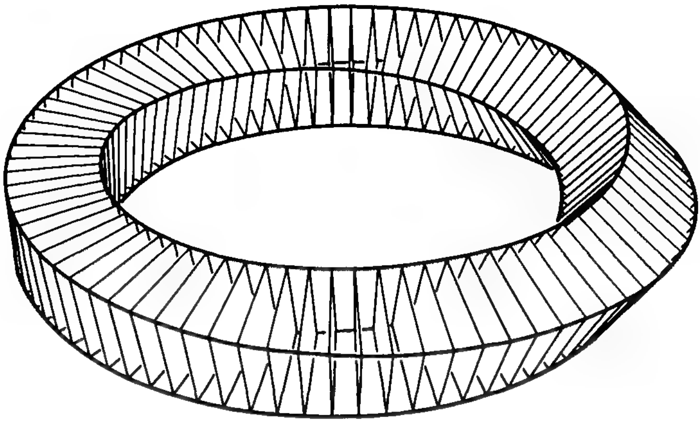


Figure 17.4 If you physically glue a cube's top to its bottom with a quarter turn you'll get a solid like this. Technically the solid has only one side which wraps around four times. But locally it has four sides. If you glue each side to its opposite you convert each square cross-section into a torus. Topologically you get the quarter-turn manifold, a circle of tori with a quarter turn in it. Note that the quarter turn is a global property of the manifold, and has nothing to do with any particular cross-section. Compare this example to the Möbius strip. Technically the Möbius strip has one edge which wraps around twice, but locally it has two edges. What surface do you get when you glue together opposite edges of the Möbius strip?

Exercise 17.3 An octagon with opposite edges glued is topologically a two-holed doughnut surface (Figure 17.6). Therefore gluing opposite side faces of an octagonal prism (Figure 17.7) produces a two-holed doughnut surface cross an interval, i.e. $(T^2 \# T^2) \times I$. Describe several ways in which you can glue the prism's top to its bottom to make a two-holed torus bundle over a circle. □

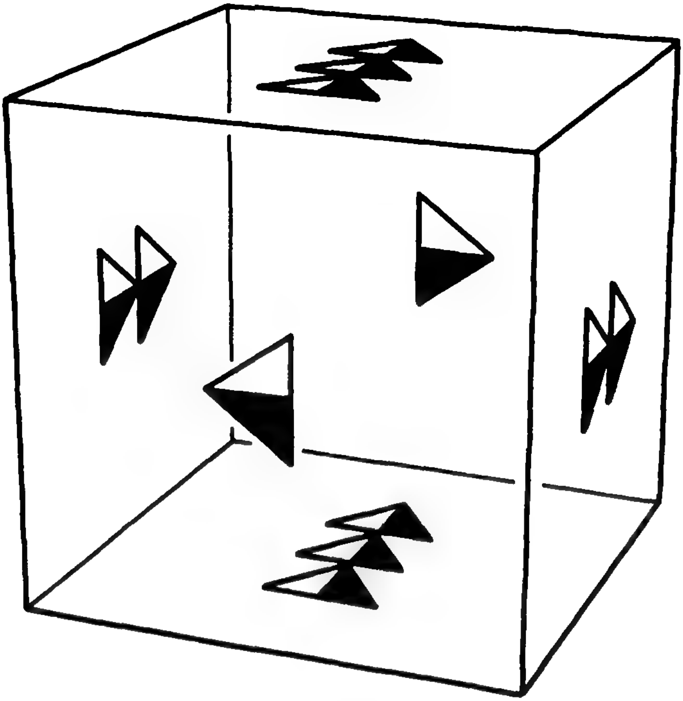


Figure 17.5 To construct $K^2 \times S^1$ you glue the cube's top to its bottom, and its left side to its right side, in the straightforward way, but you glue its front to its back with a side-to-side flip.

Each of the following exercises involves a bundle that is difficult to draw in three-dimensional space, but can be described fairly easily by gluings. If you get stuck, look up the answers.

Exercise 17.4 Name two surfaces that are circle bundles over circles. (One is a product and one isn't.)

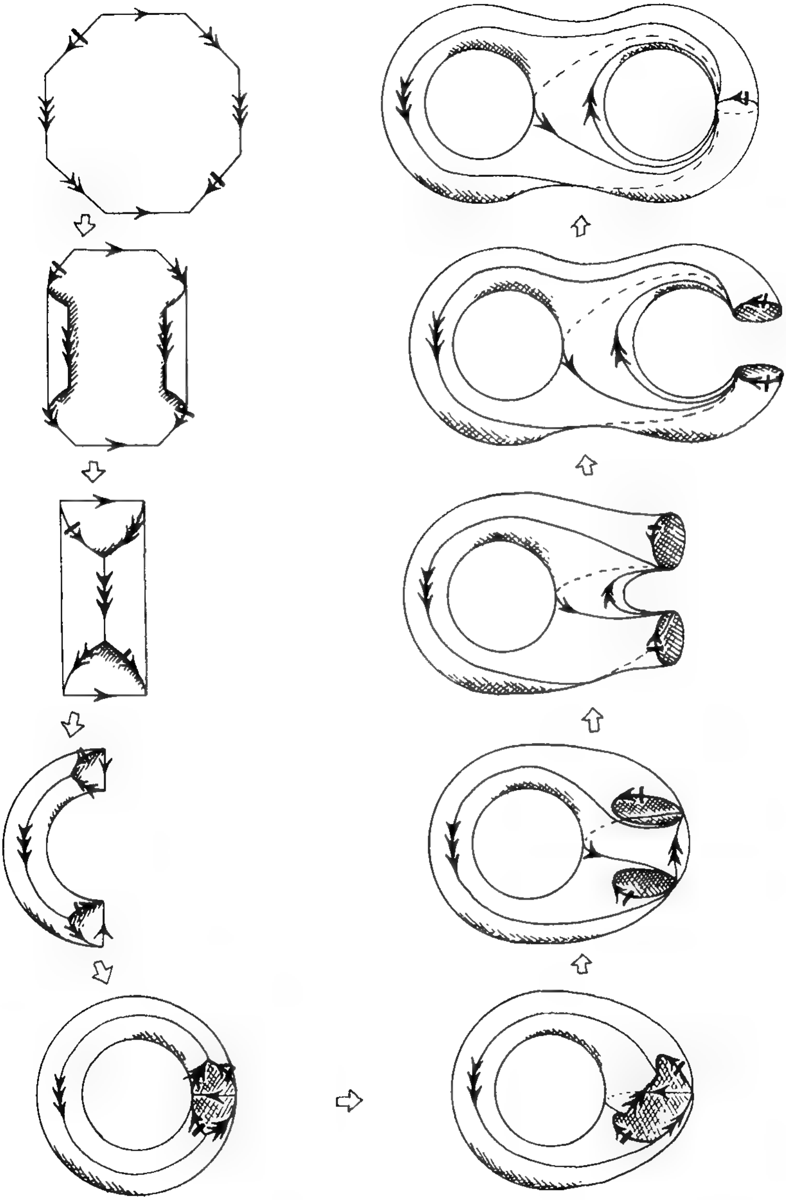


Figure 17.6 If you physically glue together opposite edges of an octagon, you'll get a two-holed doughnut surface.

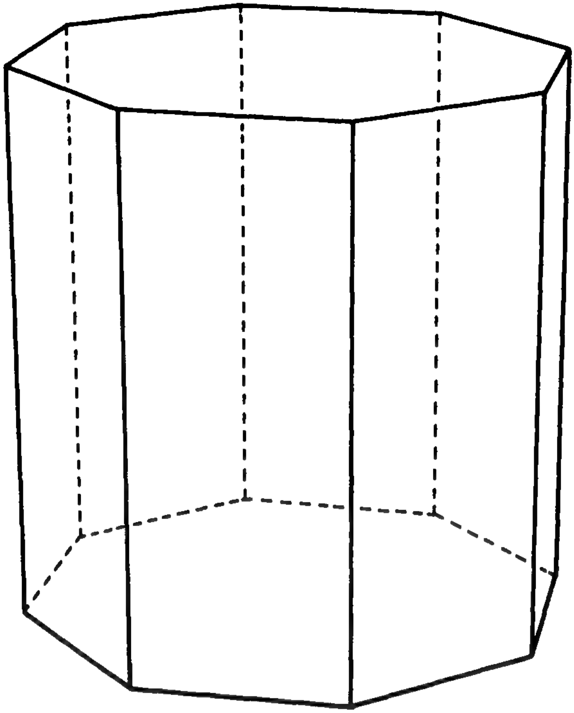


Figure 17.7 An octagonal prism with opposite sides glued is a two-holed doughnut surface cross an interval.

Exercise 17.5 A solid doughnut is topologically a disk cross a circle ($D^2 \times S^1$); its surface is topologically a torus. Describe how you would construct a *solid Klein bottle*, a disk bundle over a circle whose boundary is a Klein bottle. □

Exercise 17.6 Describe two ways in which $S^2 \times I$ can be glued up to make an S^2-bundle over S^1. One of these bundles is orientable (it's $S^2 \times S^1$). The other one is

nonorientable—in fact, it's the three-dimensional analog of a Klein bottle, so we'll denote it K^3.

What would a "solid" $S^2 \times S^1$ and a "solid" K^3 be like? I put "solid" in quotes here because $S^2 \times S^1$ and K^3 are already three-dimensional (= solid); the manifolds referred to are four-dimensional. Perhaps "hypersolid" would be more accurate? □

Exercise 17.7 Which of the flat manifolds of Chapter 7 are K^2-bundles over S^1? □

So far we've concentrated on surface bundles over circles. Now we'll switch things around and look at circle bundles over surfaces. Here's how to construct a simple example of one.

Start by packing together lots of spaghetti into the shape of a cube—the spaghetti should all be standing on end as in Figure 17.8. Thus, mathematically speaking we have a square of intervals. Glue the top of the cube to the bottom. Intrinsically this changes each piece of spaghetti from an interval into a circle, so we now have a square of circles. Glue the cube's sides together in the standard way, so that the circles, rather than being arranged in a square, are now arranged in a torus configuration. Voilà—a circle bundle over a torus! (The underlying three-manifold is, of course, just our old friend T^3.)

Exercise 17.8 Construct a circle bundle over T^2 # T^2. Start with spaghetti packed in the shape of an octag-

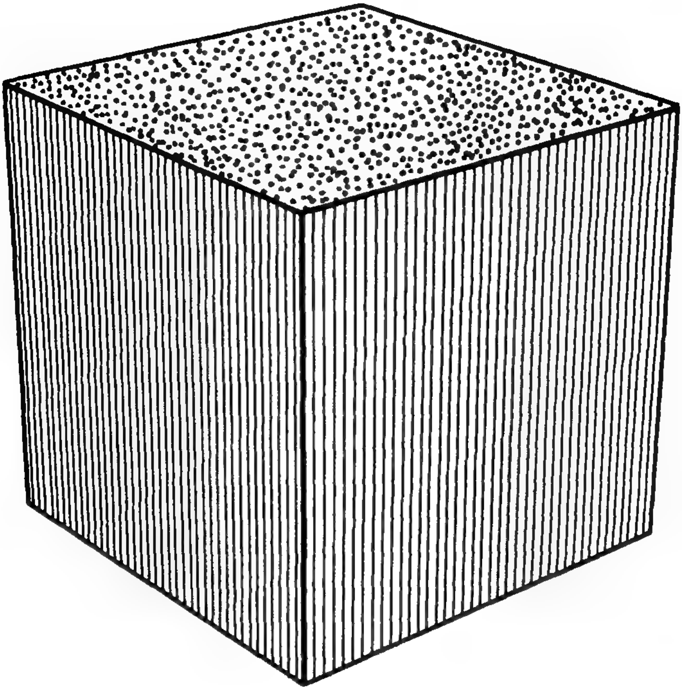


Figure 17.8 A cube filled with spaghetti standing on end.

onal prism (Figure 17.7). The resulting three-manifold will be $(T^2 \# T^2) \times S^1$. □

When we draw a picture of a circle bundle over a surface we usually won't draw in the vertical pieces of spaghetti—they tend to clutter the picture and obscure whatever else is going on. But they're there. And even if nothing is said explicitly, the top of the cube or prism will *always* be glued to the bottom, so as to convert each piece of spaghetti into a circle.

Here's a different circle bundle over a torus. Start with the cube of Figure 17.8 and glue the top to the bottom, and the left side to the right side, in the straightforward way, but glue the front to the back with a top-to-bottom flip. You'll get a torus of circles, but the circles will connect up in a strange way. (If you pass through the right face of the cube and return from the left you'll find that the circles connect up normally, but if you pass through the back face and return from the front face you'll find they connect up with a flip.) The underlying three-manifold in this case is nonorientable—it's $K^2 \times S^1$.

Exercise 17.9 Can you find examples of orientable and/or nonorientable circle bundles over the Klein bottle? (Hint: The manifolds of Figure 8.2 will come in handy.) □

Now for a really weird example! The following circle bundle over a torus, which we'll call a *twisted (three-dimensional) torus*, will play a crucial role in the theory of geometries on three-manifolds (Chapter 18).

Start with a cube as before, and, of course, glue the top to the bottom to convert the vertical intervals into circles. The sides will be glued not in any of the usual ways, but with a "shear" (Figure 17.9). By this we mean that each vertical circle gets glued to the corresponding vertical circle on the opposite side of

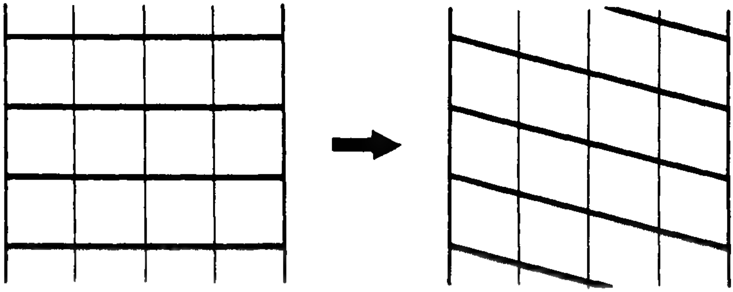


Figure 17.9 A shear.

the cube, only they get slid up or down in such a way that a horizontal line segment gets tilted. Note that for this shearing to work, it's imperative that the top of the cube be glued to the bottom. To make the twisted torus, glue opposite sides of the cube with a shear as prescribed in Figure 17.10.

This is a funny sort of gluing, and it's not at all clear that what we get is even a manifold. The following exercise deals with this (potential) problem.

Exercise 17.10 In an attempt to construct a circle bundle over a torus, the opposite faces of the cube in Figure 17.11 are glued so that the tilted lines match up. Investigate how the four vertical edges fit together. You'll discover a problem. Now check that this problem doesn't arise when the four edges come together to form the twisted torus. Conclude that the twisted torus is a bona fide three-manifold and is, in fact, a circle bundle over T^2. □

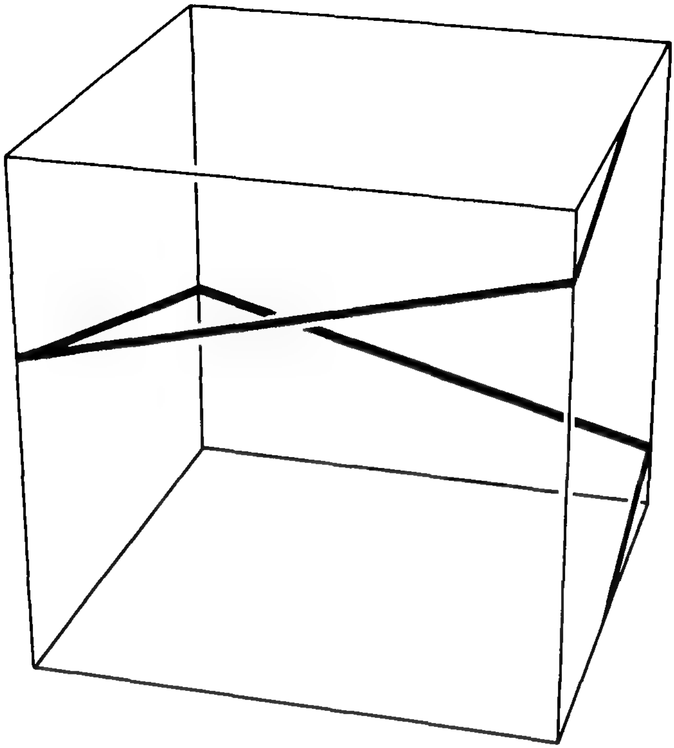


Figure 17.10 To make a twisted torus, glue the cube's opposite sides with a shear so that the tilted line segments match up. (Note: The apparently broken line segment on the right side stops being broken when the cube's top is glued to its bottom.)

So the twisted torus is a legitimate three-manifold. But is it really something new, or is it merely a distorted representation of good old T^3? If you try looking for any sort of horizontal cross-section in the twisted torus, you'll quickly become convinced that this manifold definitely *isn't* T^3.

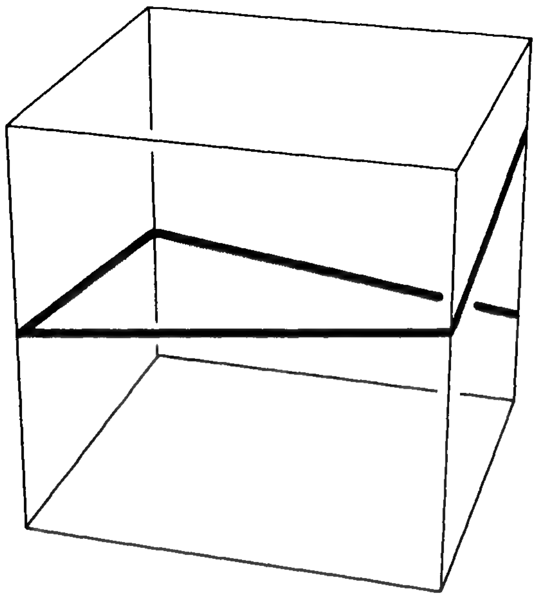


Figure 17.11 Here opposite faces are glued with only half as much shear as in the twisted torus.

Exercise 17.11 Construct a circle bundle over a torus that's twice as twisted as the twisted torus described above. □

Exercise 17.12 Construct a twisted circle bundle over $T^2 \# T^2$. □

Exercise 17.13 We constructed the twisted torus as a circle bundle over T^2. Is it also a T^2-bundle over a circle? □

18

Geometries on Three-Manifolds II

In Chapter 11 we found that every surface admits one of the three homogeneous two-dimensional geometries. The sphere and the projective plane admit elliptic geometry, the torus and the Klein bottle admit Euclidean geometry, and all other surfaces admit hyperbolic geometry. In Chapters 14, 15, and 16 we discussed the analogous three-dimensional geometries, and found some examples of three-manifolds that admit them. It turns out that five more homogeneous geometries arise in the study of closed three-manifolds. One of them is the local geometry of $S^2 \times E$.

($S^2 \times E$, $S^2 \times I$, and $S^2 \times S^1$ all have the same local geometry. It's traditional to name the geometry after $S^2 \times E$ because $S^2 \times E$ is the "biggest" manifold having it.) This geometry is homogeneous but not isotropic. It's homogeneous because it's everywhere the same. But it's not isotropic because at any given point we can distinguish some directions from others. Recall from Figure 6.10 that some cross-sections of $S^2 \times S^1$ are spheres while others are flat tori. Locally one observes that some two-dimensional slices have positive curvature while others have zero curvature (Figure 18.1). The term *sectional curvature* refers to the curvature of a two-dimensional slice of a manifold. The word "section" comes from the latin "sectio" which means "slice" (more or less). Thus $S^2 \times S^1$ has positive sectional curvature in the horizontal direction, but zero sectional curvature in any vertical slice. An isotropic geometry has the same sectional curvature in all directions; the sectional curvatures of three-dimensional elliptic geometry are all positive, those of three-dimensional Euclidean geometry are all zero, and those of three-dimensional hyperbolic geometry are all negative.

Exercise 18.1 Back in Exercise 6.7 you found that $P^2 \times S^1$ has $S^2 \times E$ geometry. Name another non-orientable manifold with this geometry. (Hint: It first appeared in Chapter 17.) □

The geometry of $H^2 \times E$ is also homogeneous but not isotropic. Like $S^2 \times E$, it has different sectional

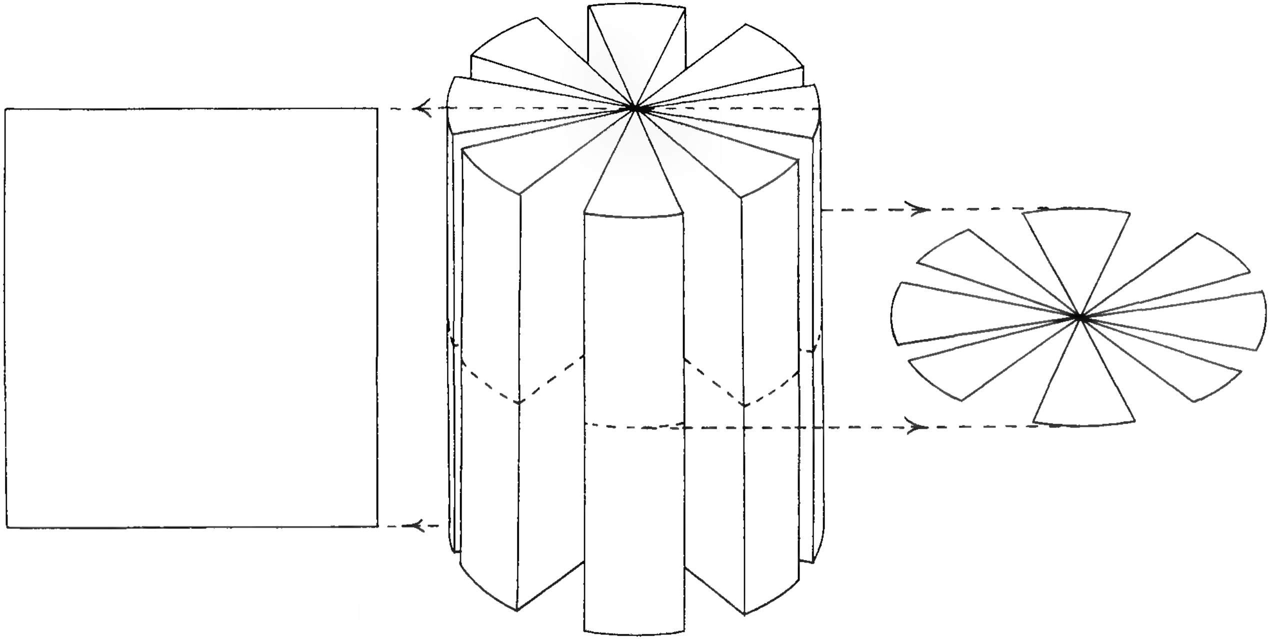


Figure 18.1 A vertical cross-section of $S^2 \times S^1$ has Euclidean geometry but a horizontal cross-section has elliptic geometry. A piece of $S^2 \times S^1$ splits vertically in Euclidean space if not allowed to stretch, but it doesn't split horizontally.

curvatures in different directions. Vertical slices have zero curvature, while horizontal slices have negative curvature. Figure 18.2 provides a rough illustration of $H^2 \times E$ geometry.

Exercise 18.2 Name several manifolds with $H^2 \times E$ geometry. □

I should mention in passing that the *only* homogeneous two-dimensional geometries are elliptic, Euclidean, and hyperbolic geometry, each of which happens to be isotropic. Geometries that are homogeneous but not isotropic occur only in manifolds of three or more dimensions.

The contemporary theory of three-manifolds deals with homogeneous geometries, without regard to isotropy. Isotropy becomes important only when one applies the mathematical theory to the study of the real universe, which appears isotropic according to current observational data. The present chapter explores homogeneous geometries in general. The cosmological applications of isotropic geometries will be treated in Chapter 19.

A couple technical points are in order before we move on to a list of the eight homogeneous geometries.

Technical Point 1: When I say "geometry" in this chapter I really mean "class of geometries." For example, the geometry of a two-sphere of radius three is, strictly speaking, different from the geometry of a

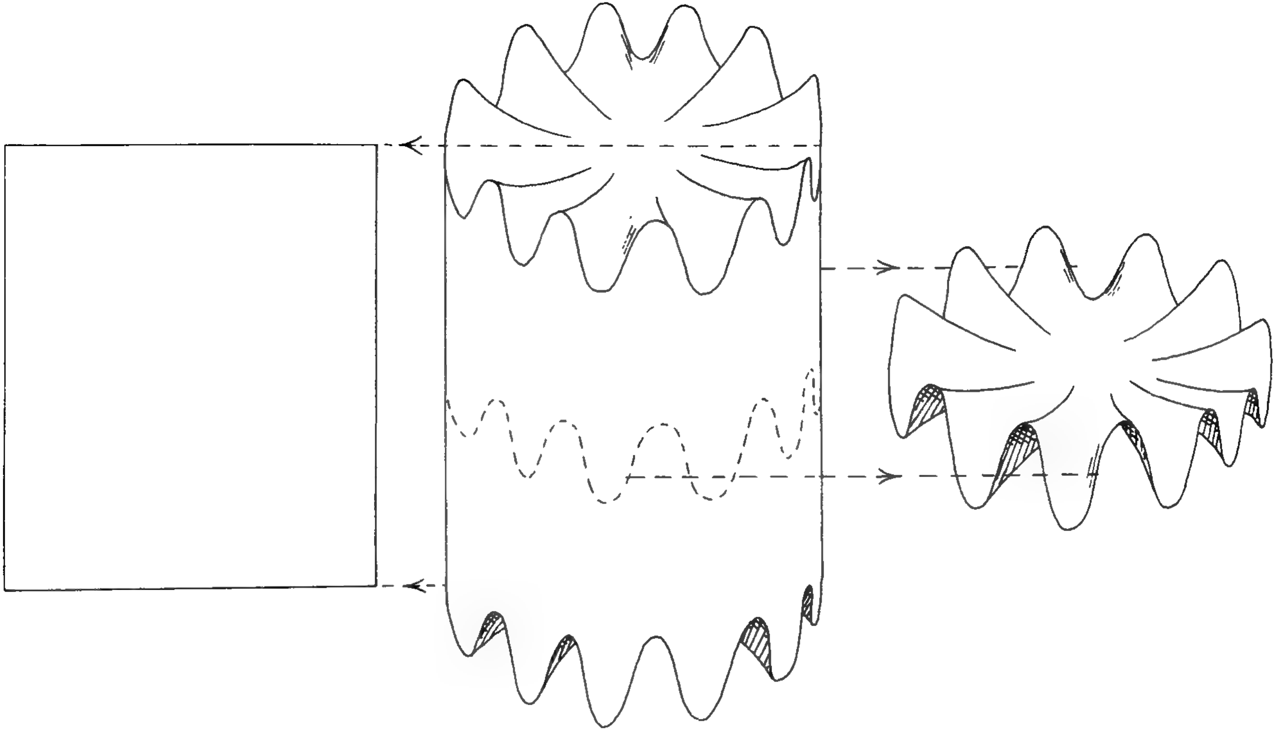


Figure 18.2 A vertical cross-section of $H^2 \times S^1$ has Euclidean geometry but a horizontal cross-section has hyperbolic geometry. By the way, the horizontal slices should really be stacked up in a different dimension than the one they wrinkle into, but this picture is the best we can do with only three dimensions available.

two-sphere of radius seventeen; yet they are similar enough that they're included in the same class of two-dimensional geometries. Three-dimensional geometries admit more variation within each class. We'll see an example of this later in the chapter.

Technical Point 2: Really there are more than eight classes of homogeneous geometries. In fact, there are infinitely many. The catch is that only eight of them occur as homogeneous geometries of *closed* three-manifolds. (As you may have noticed, this book is heavily biased towards closed manifolds!)

Here's a list of the eight homogeneous geometries, along with some sample manifolds having each one.*

(1) Elliptic Geometry

The geometry of S^3.

Sample Elliptic Manifolds: The three-sphere, projective three-space, the Poincaré dodecahedral space.

(2) Euclidean Geometry

The geometry of childhood.

Complete List of Euclidean Manifolds: Altogether there are only ten topologically different Euclidean three-manifolds! Six are orientable and four are nonorientable. The nonorientable ones are $K^2 \times S^1$, the

---

*Explore the elliptic, Euclidean, and hyperbolic examples using interactive 3-D graphics software available for free from *www.northnet.org/weeks/SoS*.

manifold of Figure 7.11, and two other K^2-bundles over S^1. The orientable Euclidean manifolds are the three-torus, the quarter turn manifold, the half turn manifold, the one-sixth turn manifold, the one-third turn manifold, and another manifold we haven't seen. The hexagonal three-torus is not included because it's topologically the same as the ordinary three-torus.

(3) Hyperbolic Geometry

See Chapter 15 for a description.

Sample Hyperbolic Manifold: Hyperbolic geometry is somewhat enigmatic. So far we've seen only one three-manifold that has it, namely the Seifert–Weber dodecahedral space. Yet research by Bill Thurston suggests that three-dimensional hyperbolic geometry is by far the most common geometry for three-manifolds, just as two-dimensional hyperbolic geometry is the most common geometry for surfaces. If hyperbolic geometry is so common, why haven't we seen more three-manifolds that have it? The reason is that the easiest manifolds to study are not the typical ones, but rather the ones with special symmetry. The first surfaces for which we found geometries were the two simplest ones, the sphere and the torus. Their homogeneous geometries (elliptic and Euclidean, respectively) are atypical precisely because these surfaces are so simple. Not until Chapter 11 did we discover that most surfaces admit hyperbolic geometry. This same phenomenon crops up in the study of three-man-

ifolds. All the simple manifolds admit atypical geometries by virtue of their simplicity. It's the typical—but less simple—manifolds that admit hyperbolic geometry.

(4) $S^2 \times E$ Geometry

This geometry was described in Chapter 6 and at the beginning of the present chapter.

Complete List of $S^2 \times E$ Manifolds: There are only four manifolds with this geometry. They are $S^2 \times S^1$, K^3, $P^2 \times S^1$, and one other manifold. This last manifold is made from $S^2 \times I$, but in this case each end is glued only to itself! Specifically, every point on an end gets glued to its antipodal point *on the same end*; the gluing resembles the gluing used to turn a ball into P^3. Note that $S^2 \times I$'s intrinsic geometry matches up nicely at the resulting "seams"; if you are confused, think about how the geometry matches up when you glue antipodal points on each end of a cylinder.

Exercise 18.3 Which of the four $S^2 \times E$ manifolds are orientable? □

(5) $H^2 \times E$ Geometry

This geometry was discussed earlier in the chapter.

Sample $H^2 \times E$ Manifolds: Here we have a little more variety than in the case of $S^2 \times E$. To begin with, any surface cross a circle will admit $H^2 \times E$ geometry, just so long as the surface *isn't* S^2, P^2, T^2, or K^2. Many

other surface bundles work too. For example, to make a (T^2 # T^2)-bundle with $H^2 \times E$ geometry, start with a two-holed doughnut cross an interval as described in Figure 17.7, and glue the top to the bottom with either a ⅛ turn, a ¼ turn, a ⅜ turn, a half turn, or one of two possible reflections.

(6) Twisted Euclidean Geometry

Figure 17.8 suggests visualizing Euclidean geometry as a bundle of vertical lines; i.e. one thinks of Euclidean space as $E^2 \times E$. Twisted Euclidean geometry may also be thought of as a bundle of vertical lines, only now the lines are bundled together in a strange new way. If you take a trip in a twisted Euclidean manifold, always traveling "horizontally" ("horizontal" means perpendicular to the vertical lines), you'll find that when you return to the line you started on you're some distance above or below your starting point! See Figure 18.3. You'll be above your starting point if you traversed a counterclockwise loop, but you'll be below it if you went clockwise. Your exact distance above or below your starting point will be proportional to the area your route encloses.

Figure 18.4 shows a jungle gym in a twisted Euclidean manifold. All the bars have the same length, and they all meet at right angles. Some bars *appear* tilted because the artist had to distort the twisted Euclidean jungle gym to draw it in ordinary Euclidean space.

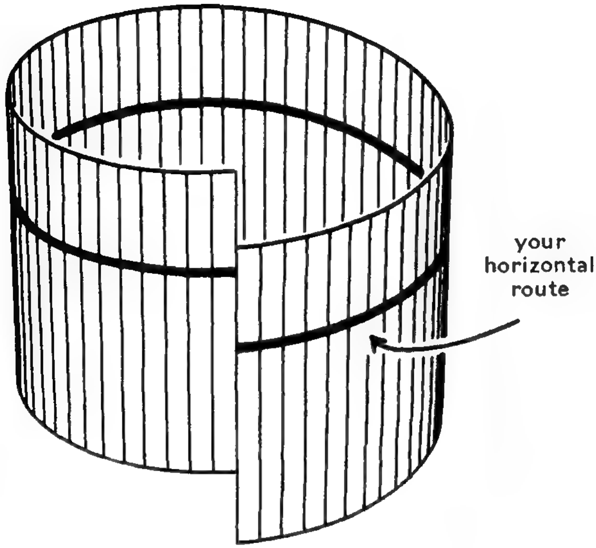


Figure 18.3 If you travel horizontally in a twisted Euclidean manifold you'll come back to a point above or below where you started.

Sample Twisted Euclidean Manifolds: You can give the twisted torus of Chapter 17 a twisted Euclidean geometry. First put a twisted Euclidean geometry on the cube of Figure 17.10. With this new geometry the previously tilted lines on the sides of the cube become intrinsically horizontal. When you glue opposite sides of the cube with a shear, you glue the vertical lines on one side to the vertical lines on the other, and the horizontal lines on one side to the horizontal lines on the other. Thus, in terms of the twisted Euclidean geometry you're gluing one side rigidly to the other, with no "shearing" or other abnormalities. The corners and

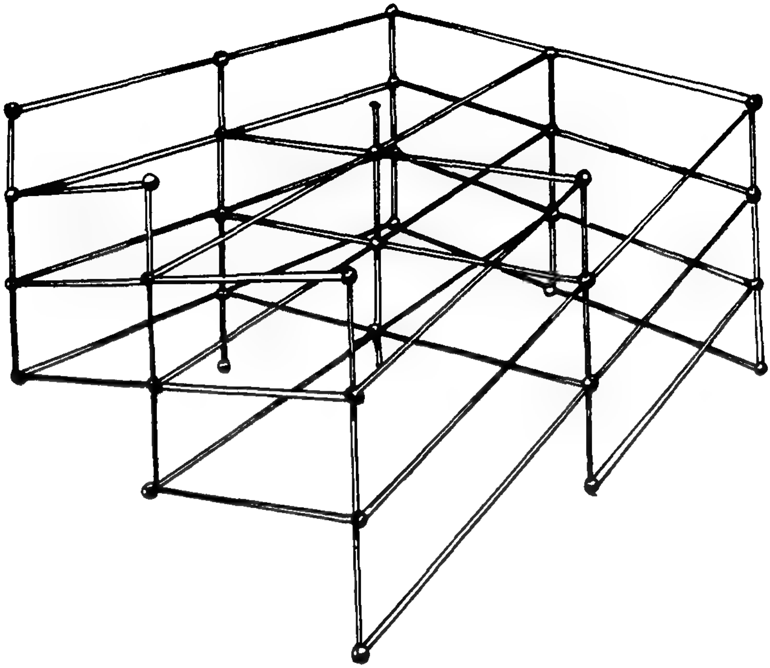


Figure 18.4 A jungle gym in a twisted Euclidean manifold. (Drawing by Bill Thurston.)

edges still fit fine, so you've given the twisted torus a twisted Euclidean geometry.

The doubly twisted torus of Exercise 17.11 also admits twisted Euclidean geometry. In fact *every* circle bundle over a torus or Klein bottle admits either Euclidean geometry or twisted Euclidean geometry.

Exercise 18.4 Can a nonorientable three-manifold have a twisted Euclidean geometry? □

(7) Twisted $H^2 \times E$ Geometry

This geometry bears the same relation to standard $H^2 \times E$ geometry as twisted Euclidean geometry does to standard Euclidean geometry. Specifically, one thinks of both the twisted and standard $H^2 \times E$ geometries as vertical line bundles over H^2. In the twisted case the lines connect up as in Figure 18.3.

Sample Twisted $H^2 \times E$ Manifolds: Any circle bundle over any surface except S^2, P^2, T^2, or K^2 admits either standard or twisted $H^2 \times E$ geometry.

Exercise 18.5 Give explicit instructions for constructing a sample twisted $H^2 \times E$ manifold. □

(Twisted $S^2 \times E$ Geometry): Amazingly enough, if you put the right amount of twist into $S^2 \times E$ you'll get elliptic geometry! The amount of twist is right when a traveler traveling horizontally around a region of area α returns to a point α units below where she started (here I assume that $S^2 \times E$ is a bundle over a *unit* two-sphere). Even when the amount of twist is wrong this geometry is still classified as elliptic geometry, as per Technical Note 1, because its group of symmetries is contained in the group of symmetries of the three-sphere. Note, though, that you can always adjust the twist to the right value by stretching or compressing the vertical lines.

(8) Solve Geometry

This is the real weirdo. Unlike the previous geometries, solve geometry isn't even rotationally symmet-

ric. I don't know any good intrinsic way to understand it. (The name "solve" geometry has to do with "solvable Lie groups.")

Sample Solve Manifolds: Most torus bundles over S^1 admit solve geometry. (None of the ones we've seen in this book do, though, because none of them distort the geometry of the two-dimensional cross-section.)

It would be nice if every three-manifold admitted one of the above homogeneous geometries. Alas, this is not the case. For example, a connected sum of two three-manifolds never admits a homogeneous geometry (unless either one of the original manifolds is S^3, or both original manifolds are P^3). Fortunately the situation isn't quite as bad as it sounds. Thurston's work suggests that most three-manifolds admit hyperbolic geometry, and those that don't either admit one of the other seven homogeneous geometries, or can be cut into pieces that admit homogeneous geometries. (In this context one "cuts a manifold into pieces" by cutting it along spheres, projective planes, tori, and/or Klein bottles.) This view of three-manifolds has not yet been proven correct, but there is a good deal of evidence in its favor. For a summary of the evidence, see Thurston's article *Three dimensional manifolds, Kleinian groups and hyperbolic geometry* (Bulletin of the AMS 6(1982), pp. 357–381). An interesting corollary of Thurston's ideas is that a "randomly chosen" three-manifold is unlikely to be a connected sum.

This Page Intentionally Left Blank

Part IV

The Universe

This Page Intentionally Left Blank

19

The Universe

The universe has existed for only 10 or 15 billion years. This chapter discusses the beginning of the universe (the big bang), the ensuing expansion, and the relationship between the shape of the universe and the matter it contains. The chapter is organized around the following questions:

1. What do we know about the universe?
2. In what sense is the universe expanding?
3. How is the density of matter related to the curvature of space?

4. Is the universe closed or open? In other words, is space finite or infinite?
5. What came before the big bang?

Question 1. What do we know about the universe?

We live on Earth, which is one of approximately nine planets orbiting our sun.* The sun is grouped together with about 100 billion other stars in our galaxy the Milky Way. The Milky Way is shaped like a giant disk 50 to 100 thousand light-years across,† with us about 25 thousand light-years from the center. If you go out on a clear, moonless night the stars of our galaxy will appear as a splotchy white band running across the sky, hence the name Milky Way. The other stars you normally see are all in our galaxy too.

There are lots of other galaxies in the universe—billions of them at least. Some are off by themselves, but most lie in clusters of anywhere from a few to a few thousand galaxies each. On average, galaxies are distributed in space roughly like dimes spaced a meter apart.

A striking feature of these other galaxies is that they are moving away from us! This is *not* because the other galaxies are moving through space. Rather, the

---

*By modern standards Pluto wouldn't qualify as a planet, but tradition is strong so Pluto is likely to retain its planet status for the foreseeable future.

†A light-year is the distance light can travel in one year. By way of comparison, the sun's light takes about 8 minutes to reach us, so the sun is 8 light-minutes away. Similarly, the moon is about 1 light-second away, and San Diego is about 0.02 light-seconds from Boston.

galaxies are more-or-less still relative to space, and it's space itself that is expanding, carrying the galaxies along with it. Observations show that space is currently expanding at a rate of about 7% per billion years. In other words, if the universe were to continue expanding at its present rate, after a billion years all cosmic distances would be stretched by 7%.

Exercise 19.1 Galaxy A and galaxy B presently lie 15 billion light years away from each other. If the universe expands at a constant rate of 7% per billion years, how far apart will galaxies A and B be a billion years from now? How fast is galaxy B moving away from galaxy A? Is this slower or faster than the speed of light? □

The concept of an expanding universe has an interesting history. When Einstein in 1917 first applied his geometric theory of gravity (his famous theory of general relativity) to the universe as a whole, he found his equations inconsistent with a universe of constant size. Surprised and perplexed, he introduced a "cosmological constant" Λ into his equation as a fudge factor to make his constant-size universe work. A few years later, in 1922, Alexander Friedmann conceived the idea of an expanding universe. To his delight, he found that Einstein's original equations worked fine in an expanding universe, with no need for a fudge factor. Nevertheless the idea of an unchanging universe was so ingrained in western thinking that not even Einstein could accept Friedmann's work: Ein-

stein regarded Friedmann's model of an expanding universe as a mere mathematical curiosity without physical significance.

Fortunately experimental support was quick in coming. During the 1920s the work of many astronomers contributed to the discovery that other galaxies are receding from us. The exciting conclusion, that a galaxy's rate of recession is roughly proportional to the distance from us to the galaxy, was exactly what one would expect in an expanding universe. History has assigned the credit for this conclusion to Edwin Hubble and recorded the date as 1929. In reality Georges Lemaître had come to the same conclusion two years earlier, in 1927, using essentially the same data and computing the same value for the rate of expansion. Indeed there is a bit of a scandal here. Lemaître's original paper was in French and not widely read. When Eddington translated Lemaître's paper into English in 1931, he completely omitted the paragraph in which Lemaître computed the expansion rate, and even took care to excise the expansion rate from a subsequent equation in which it appeared! Nevertheless, to this day we call the expansion rate the Hubble constant and denote it by the letter H, that is, $H = 7\%$ per billion years.

The fact that space is expanding means that in the past it must have been smaller. If we look far enough back in time, space had zero size. This was the big bang, the birth of the universe. How long ago was the big bang? To get a rough idea, assume space

has been expanding at a constant rate of 7% per billion years.* To get back to zero size would require (100%)/(7%/billion years) ≈ 15 billion years. We'll take a more careful look at the big bang later in this chapter, and see direct physical evidence of it in Chapter 22.

Observation shows that at least the visible portion of the universe is both homogeneous and isotropic.

Homogeneity

This means that any two regions of the universe are basically alike. Of course, we have to look on a sufficiently large scale, so that "local" fluctuations in the number of galaxies get averaged out. The situation is analogous to saying that a roomful of air is homogeneous, even though one cubic microcentimeter might contain 17 molecules of nitrogen and 4 of oxygen, while a different cubic microcentimeter might contain 8 of nitrogen and 11 of oxygen.

Isotropy

This means that no matter where you are in the universe, things look basically the same in all directions. An isotropic universe is of necessity homogeneous—to see that conditions must be the same at any two locations A and B, note that the universe is isotropic

---

*By this we mean that during each billion year period the universe grew by 7% of today's size, not 7% of the size it was during that period. In other words, assume the universe grew at a constant linear rate, not a constant exponential rate.

about the point lying halfway between A and B. The visible portion of the real universe is known to be isotropic because the number of galaxies is roughly the same in all directions, the expansion rate (the Hubble constant) is the same in all directions, and, best of all, the cosmic microwave background radiation (Chapter 22) is the same in all directions to the precision of a few parts in 10^5.

Even though the *visible* portion of the universe is approximately homogeneous and isotropic, the universe *as a whole* could well be inhomogeneous, with curvature that varies gradually from one part of space to another. In other words, the visible universe, as vast as it is, might be only a tiny portion of the whole universe, too small to reveal large-scale variations in curvature. Even though this hypothesis is completely plausible, it holds little interest for topologists because in such a huge universe we would not be able to directly observe the topology of space. It would be as if the biplane pilot in Figure 7.1 could see no further than one plane length in any direction—she wouldn't be able to see that her universe is a torus.

If we make the (still unconfirmed!) assumption that the universe is small enough that we may observe its topology directly (for example by observing multiple images of the same galaxy, just as the biplane pilot in Figure 7.1 observes multiple images of the same biplane), then the observable universe *is* the whole universe, and so the well-tested homogeneity

and isotropy of the visible universe imply that space as a whole is homogeneous and isotropic. For this reason all studies of the possible topologies of space, and all efforts to detect the topology experimentally, have focused on homogeneous, isotropic three-manifolds. Fortunately there are only three homogeneous, isotropic local geometries for us to consider, namely elliptic geometry (Chapter 9), Euclidean geometry, and hyperbolic geometry (Chapter 10). Of course there are many different three-manifolds having each geometry, and thus many possible global topologies for the universe. For example, the Poincaré dodecahedral space has locally elliptic geometry, the three-torus has locally Euclidean geometry, and the Seifert–Weber space has locally hyperbolic geometry.

Question 2. In what sense is the universe expanding?

Figure 19.1 illustrates an *incorrect* answer to this question. The big bang was *not* like a giant firecracker exploding into an already existing space. The big bang had no center.

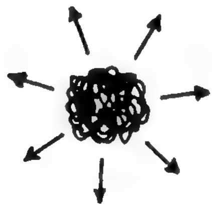
Figure 19.2 illustrates the correct answer to the question. Space itself was very small right after the big bang, and didn't even exist before it! Note that each galaxy sees neighboring galaxies receding from it, just as Lemaître, Hubble, and their colleagues observed in the real universe in the 1920s.

The expansion of a three-dimensional torus universe is shown in Figure 19.3. Naturally this idea applies to a universe based on any three-manifold.

An Incorrect Picture of the Big Bang

(1) At the moment of the big bang, all matter starts out at a single point in space.

(2) It goes flying off into space in all directions,



(3) and eventually forms galaxies which continue to move further out into space.

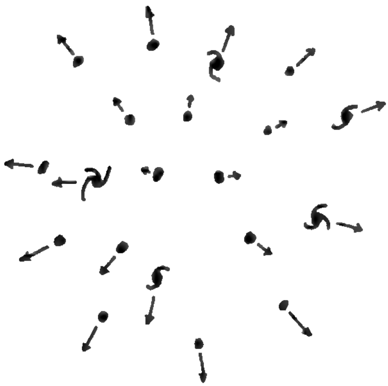


Figure 19.1

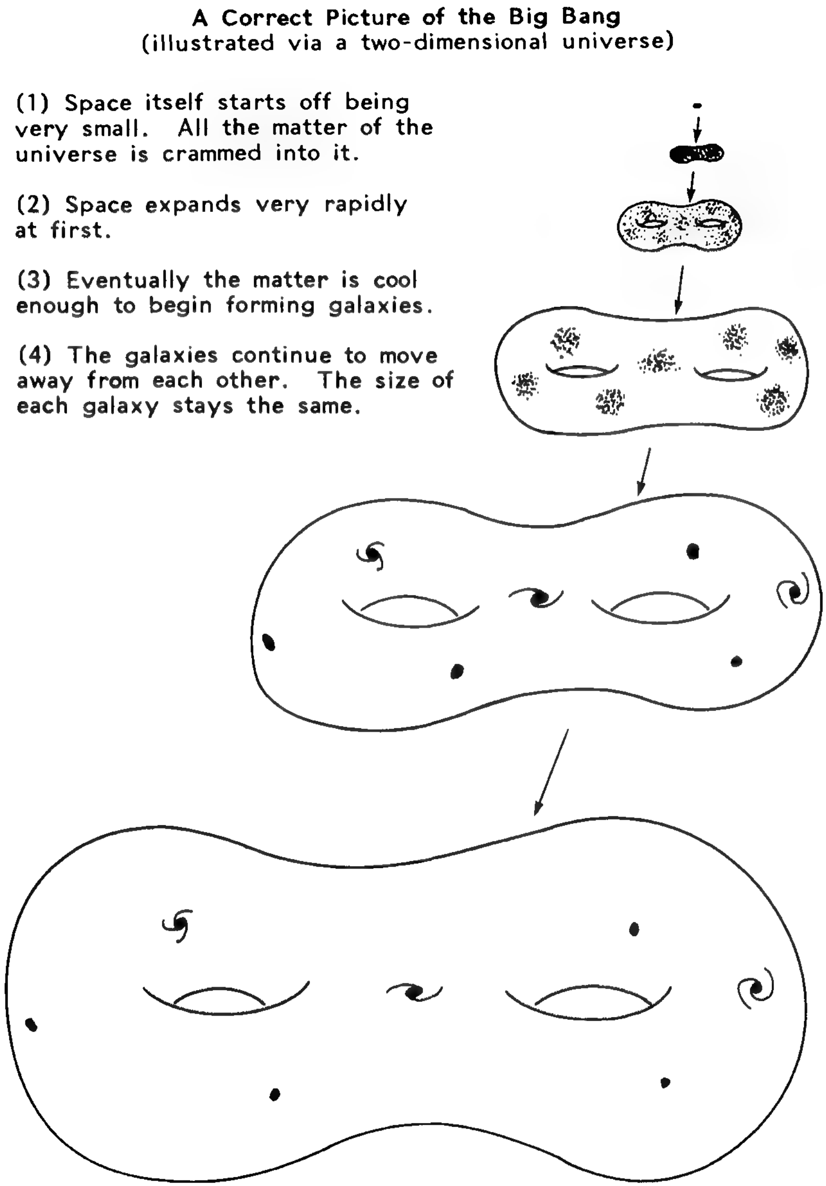
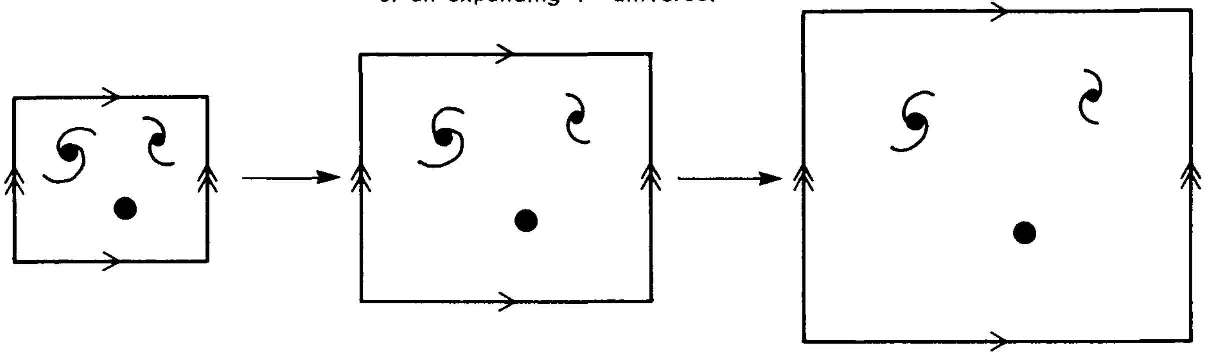

A Correct Picture of the Big Bang
(illustrated via a two-dimensional universe)
(1) Space itself starts off being very small. All the matter of the universe is crammed into it.
(2) Space expands very rapidly at first.
(3) Eventually the matter is cool enough to begin forming galaxies.
(4) The galaxies continue to move away from each other. The size of each galaxy stays the same.

Figure 19.2

For comparison, here's a Flatlander's drawing of an expanding T^2-universe:



Now here's a Homo Sapien's drawing of an expanding T^3-universe:

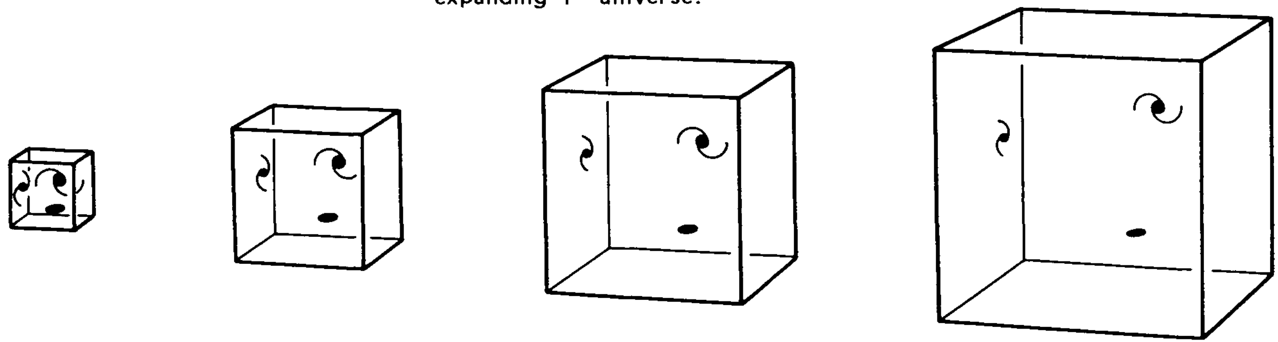


Figure 19.3

One question that often comes up is that if everything is expanding—houses, people, atoms, metersticks, *everything*—then how can we tell that things have changed at all? The answer is that not everything is expanding. Houses, people, atoms, and metersticks are not expanding. Planets, stars, and even galaxies are not expanding. Space *is* expanding, and so is the distance between galaxies, but that's about it.

Another question that comes up is that if the universe has infinite volume (i.e., if it's an open three-manifold), then how can it expand and get any bigger? The answer is that its total volume doesn't increase, but space does still stretch out, and the distances between galaxies do still increase, just as they would in a closed universe. The expansion of an infinite universe is *locally* identical to the expansion of a finite one. You can, for example, reinterpret Figure 19.3 as an expanding chunk of space in an infinite universe. The difficulties that arise when contemplating the total volume of an infinite expanding universe are difficulties with the concept of infinity, not difficulties with the behavior of the universe.

Question 3. How is the density of matter related to the curvature of space?

When Friedmann applied Einstein's theory of general relativity to the idea of an expanding universe, he found a relationship between the density of matter, the rate of expansion, and the curvature of space. Spe-

cifically, he found that in a universe with elliptic geometry (such as the three-sphere or the Poincaré dodecahedral space) the average density ρ of matter and energy must be greater than a certain minimum given by the formula

$$\rho > \frac{3}{8\pi G} H^2 \quad \text{(elliptic universe)}$$

where H is the Hubble constant and G is the constant from Newton's law of gravitation $F = GmM/r^2$. In a hyperbolic universe (such as the Seifert–Weber space) the density ρ of matter and energy must be less than that same critical amount

$$\rho < \frac{3}{8\pi G} H^2 \quad \text{(hyperbolic universe)}$$

and in a flat universe (such as the three-torus or quarter turn space) the density ρ must exactly equal the critical amount

$$\rho = \frac{3}{8\pi G} H^2 \quad \text{(flat universe)}$$

This is an exciting discovery, because both the density ρ and the Hubble constant H can be measured experimentally, thus allowing us to deduce the curvature of space. Many such measurements have been made, including familiar matter such as stars as well as the poorly understood "dark matter" contained in galaxies. Such studies consistently find that the total mass density ρ is only about 30% of the critical amount, implying that the universe is hyperbolic.

On that basis, researchers studying the topology of space considered mainly hyperbolic models. But in 1998 the situation began to change dramatically. New data (coming from studies of distant supernovas and the cosmic microwave background radiation) made a strong case that the visible universe is not hyperbolic, but flat. At first glance this seems to contradict the earlier studies that found the density ρ to be only 30% of the critical amount, not 100%. Happily there is no contradiction. The resolution lies in the fine print. The earlier studies (the 30% result) measured the density of matter in the universe. The newer studies (the 100% result) indirectly measure the density of matter *and energy*. The conclusion, then, is that while matter contributes 30% of the critical density, some sort of mysterious vacuum energy contributes the remaining 70%. As of Autumn 2001 little is known about the vacuum energy. Indeed at this point it's fair to say that the term "vacuum energy" is just a hollow label used to refer to a concept about which we know nothing. Nevertheless, the vacuum energy will surely play a key role as our understanding of the universe develops over the 21st century.

Question 4. Is the universe closed or open? In other words, is space finite or infinite?

Put briefly, we don't know.

We do know, however, that if the universe has an elliptic geometry then it must be closed. If, on the other hand, the universe has a flat or hyperbolic ge-

ometry, then it can be either closed or open. Table 19.1 gives some sample topologies for the different types of universes.

When the first edition of this book appeared in 1985, many cosmologists were completely unaware of closed manifolds with flat or hyperbolic geometry. The situation has improved greatly since then, but you might still find textbooks stating incorrectly that a flat or hyperbolic universe must be infinite. The terminology that grew up around this misconception is particularly unfortunate: in the cosmological literature "closed" is used to mean "elliptic" (= positively curved), while "open" is used to mean "hyperbolic" (= negatively curved) and "critical" is used to mean "flat" (= zero curvature). This terminology precludes the very mention of a closed flat universe or a closed hyperbolic one.

Question 5. What came before the big bang?

Sagredo: What came before the big bang?
Salviati: Nothing did.
Sagredo: You mean space was entirely empty then?
Salviati: No, space didn't even exist!
Sagredo: Oh, I see: at times before the big bang there simply was no space. What a curious thought.
Salviati: It's worse than that: "before" the big bang there wasn't any time either!
Sagredo: What? No time?! Even if there was no matter and no space, surely there would have been time.
Salviati: Allow me to draw you some pictures. They'll be *spacetime diagrams* somewhat like the one in Figure 13.8, only these will be pic-

Table 19.1 Possible Global Topologies for the Universe

	Elliptic geometry	Euclidean geometry	Hyperbolic geometry
Closed	S^3, P^3, Poincaré dodecahedral space	T^3, ¼ turn manifold, ⅙ turn manifold	Seifert–Weber space, "most" closed 3-manifolds (see Chapter 18)
Open	None	E^3, $T^2 \times E$, $E^2 \times S^1$	H^3 (there are other possibilities as well)

tures of a one-dimensional circular universe rather than a two-dimensional planar one. If this circular universe always stays the same size, then its spacetime diagram will be a cylinder, as in Figure 19.4 (left). On the other hand, if it's an expanding circular universe, then its spacetime diagram will be a cone, as in Figure 19.4 (right).

The import of these pictures is that space and time have been wrapped into a unified *spacetime*. So where there's no space there's no time, and vice versa. I might add that Einstein's relativity not only permits us, but actually forces us to think of space and time in this way.

Sagredo: I see. Spacetime includes *all* matter, *all* space, *and* all time—in short, *all* of physical reality.

Salviati: Right.

Sagredo: And all of physical reality comes *after* the big bang.

Salviati: Exactly.

Sagredo: So what *caused* the big bang?

Salviati: Nothing did.

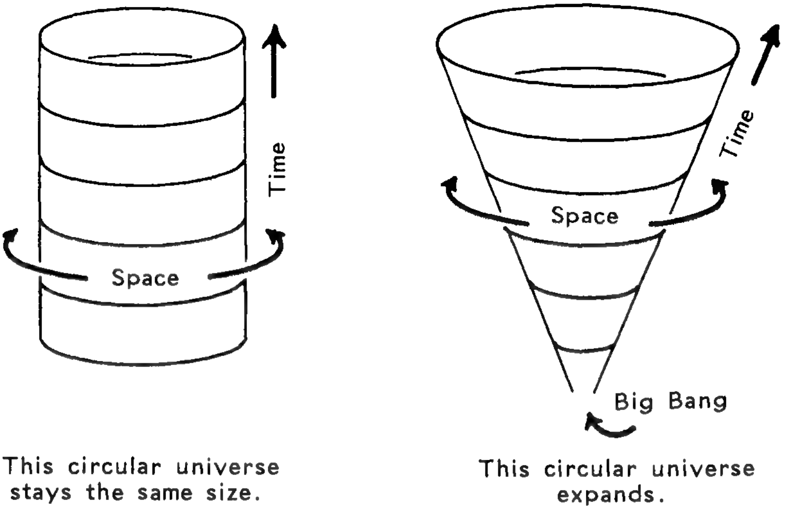


Figure 19.4 Spacetime diagrams for static and expanding circular universes.

Sagredo: What?! Something must have caused it!

Salviati: Physical events are caused by other physical events, but the big bang wasn't an ordinary physical event, and doesn't demand a cause.

Sagredo: But the big bang appears in the spacetime diagram of Figure 19.4 (right), so doesn't that make it a physical event?

Salviati: I'm sorry, I should have explained that earlier. The spacetime diagram you refer to is meant to be a cone with no vertex. Thus, strictly speaking, there is no big bang in that diagram, there are only events following it. Every event is caused by earlier events, yet there is no earliest event in the spacetime diagram in the same way that there is no smallest positive number. Of course we may hope that future generations will achieve a deeper understanding of the

big bang and fill in the missing part of the picture. For example, quantum cosmologists are already exploring theories in which the cone is capped off with a tiny hemisphere.

Sagredo: So what caused spacetime?

Salviati: Now that's a difficult question! I myself have no idea. In fact, I'm not even sure that the question is a meaningful one, at least not in the ordinary sense of the word, that one event in spacetime causes another event in spacetime. But if you want to postulate a god as the cause of the universe, then according to relativity theory you should imagine Him to have caused spacetime as a whole, rather than just the big bang.

Sagredo: You mean He created the past and the future at the same time?

Salviati: That's the idea, although "at the same time" is a misleading choice of words. Just as God doesn't have a location in space, neither does He experience time. A god would have to be outside spacetime, and being outside spacetime means being outside both space *and time*.

Sagredo: That's certainly a switch from the traditional view, in which God is the creator of space, but nevertheless lives in time just as we do.

If you don't mind, Salviati, I'd like to leave these theological issues aside and ask you one more question.

Salviati: Please do.

Sagredo: How do you know your description of the big bang is correct?

Salviati: I don't! The *nature* of the big bang is very much a matter of speculation, even though its *existence* is supported by overwhelming

evidence, such as the cosmic microwave background radiation described in Chapter 22.

Sagredo: Could you tell me about some alternative descriptions of the big bang?

Salviati: Certainly. In Figure 19.5 I've sketched a number of spacetime diagrams, each of which gives a different picture of the universe. The first diagram depicts a circular universe that expands to some maximum size and then recollapses to a big crunch, while the second depicts a circular universe that expands forever. Diagrams #3 and #4 depict oscillating universes—in one case the universe collapses to a point at the end of each cycle, and in the other it "bounces" before it gets that far. Finally, diagrams #5 and #6 depict cyclic universes in which the big crunch is the cause of the big bang.

Sagredo: Don't the laws of physics tell us which of these models is correct?

Salviati: Unfortunately not. Strange things happened in the first zillionth of a second after the big bang.* Temperatures, pressures, and densities were enormous, perhaps unboundedly so. Under these conditions gravity takes on a quantum nature, but we have as yet no quantum theory of gravity. Thus we can't say with any certainty what the universe was like at times close to the Big Bang, and so we must content outselves with idle speculation.

Sagredo: So tell me, which picture do you advocate?

Salviati: I prefer diagrams #1 and #2 because they are the simplest. Current observational evidence suggests that the expansion of the universe is accelerating, which would imply

*In this case a zillionth is approximately 10 ^{43}.

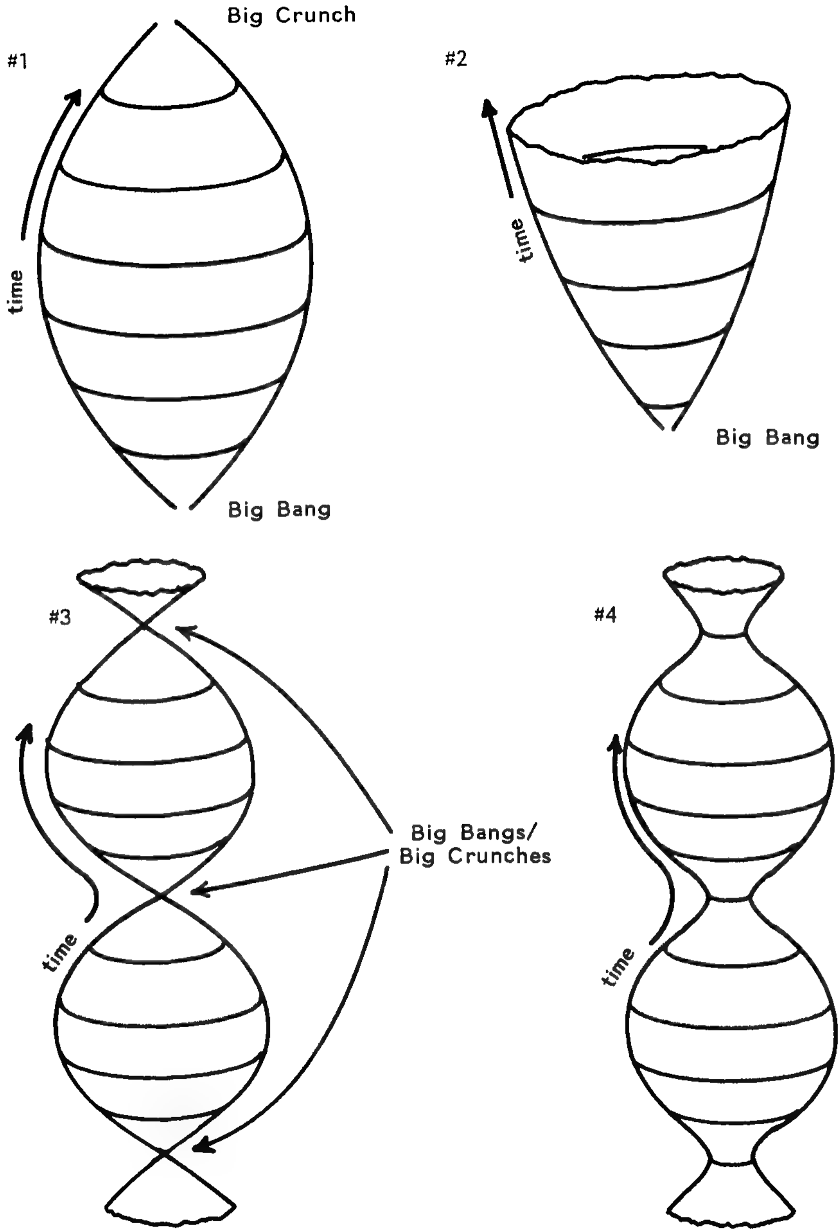


Figure 19.5 Alternative interpretations of the big bang.

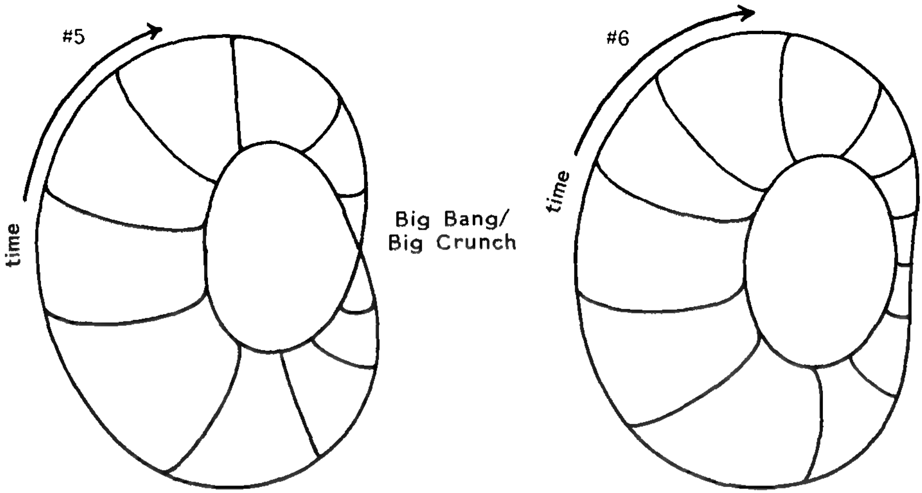


Figure 19.5 Continued.

that diagram #2 is more accurate than diagram #1.

The chief appeal of the remaining diagrams is that one needn't part with the cherished notion that time has neither a beginning nor an end. But these other pictures introduce other complications—namely one must either interpret the point representing the big bang (diagrams #3 and #5) or explain why the universe bounces (diagrams #4 and #6). I myself would rather accept the fact that our past may be finite.

20

The History of Space

The ancient Greeks had varying opinions on the nature of space. Leucippus (ca. 480 B.C.–ca. 420 B.C.) and Democritus (ca. 460 B.C.–ca. 370 B.C.) imagined an infinite universe, but Aristotle (ca. 384 B.C.–ca. 322 B.C.) envisioned the universe as a finite ball with the Earth at its center and a spherical boundary. Aristotle's views prevailed, and went largely unquestioned in Western society for 2000 years, although in China and perhaps elsewhere there was a belief in an infinite universe.

The invention of the telescope in 1608 led to new observations undermining the geocentric picture, and rekindling interest in an infinite universe. Some people were uncomfortable with an infinite universe, but didn't like Aristotle's boundary either. What would lie beyond the boundary? In his 1854 Habilitationschrift, Riemann proposed the three-sphere as a model of the universe. Recall from Chapter 14 that the three-sphere is the three-dimensional surface of a four-dimensional ball. It's a finite universe, yet has no troublesome boundary.

In 1890 Klein found a much more general solution, namely the idea of a multiconnected universe. The simplest multiconnected three-manifold is the three-torus. Roughly speaking, a three-manifold is called multiconnected if you see multiple images of yourself, as explained in Figures 7.1 and 7.2, and further illustrated in Figures 7.3 through 7.12. All closed three-manifolds in this book, except the three-sphere, are multiconnected. The famous Poincaré conjecture claims that no other exceptions are possible, but this conjecture remains unproved.

Astronomers initially took an interest in the idea of a multiconnected space. As early as 1900 Karl Schwarzschild presented the three-torus in a postscript to an article in the Vierteljahrschrift der Astronomischen Gesellschaft, challenging his readers to

> . . . imagine that as a result of enormously extended astronomical experience, the entire Universe consists of countless identical copies of our Milky Way, that the

> infinite space can be partitioned into cubes each containing an exactly identical copy of our Milky Way. Would we really cling on to the assumption of infinitely many identical repetitions of the same world? In order to see how absurd this is consider the implication that we ourselves as observing subjects would have to be present in infinitely many copies. We would be much happier with the view that these repetitions are illusory, that in reality space has peculiar connection properties so that if we leave any one cube through a side, then we immediately reenter it through the opposite side. The space that we have posited here is nothing other than the simplest Clifford–Klein space [*the three-torus*], a finite space with Euclidean geometry. One recognizes immediately the sole condition that astronomical experience imposes on this Clifford–Klein space: because visible repetitions of the Milky Way have not yet been observed, the volume of the space must be much greater than the volume we ascribe to the Milky Way on the basis of Euclidean Geometry. The other simple Clifford–Klein spaces can be dealt with briefly because their mathematical study is incomplete.* They all arise in the same way through apparent identical copies of the same world, be it now in a Euclidean, elliptic or hyperbolic space, and our experience imposes the condition that their volume must be bigger than that of the visible star system.†

When Einstein applied his newly conceived geometrical explanation of gravity (his famous theory of general relativity) to the questions of cosmology in 1917,

---

*W. Threlfall and H. Seifert classified spherical manifolds by 1930 and W. Hantzsche and H. Wendt classified flat manifolds by 1935, but hyperbolic manifolds remain unclassified to this day.

†K. Schwarzschild, "On the permissible curvature of space," Vierteljahrschrift d. Astronom. Gesellschaft 35 (1900) 337–347; translated into English by John and Mary Stewart, Class. Quantum Grav. 15 (1998) 2539–2544. Used with permission.

he chose Riemann's three-sphere as his model of space. Einstein's colleague de Sitter, however, was quick to point out, still in 1917, that Einstein's equations for a spherical universe applied equally well to the multiconnected projective three-space (see the last section of Chapter 14). Alexander Friedmann took the even more remarkable step, in his 1924 paper on the possibility of a hyperbolic universe, of pointing out that his equations applied in principle not only to infinite hyperbolic space but also to all closed (multiconnected) hyperbolic three-manifolds, even though not a single example was known at the time! Nevertheless, Einstein continued to prefer the three-sphere because of its simplicity. Indeed, Einstein's primary motivation for modeling the universe as a spherical manifold instead of a flat or hyperbolic one was that the three-sphere alone offered the possibility of a universe both finite and simply connected.

During the middle years of the twentieth century, cosmologists lost interest in the question of the topology of the universe. Their lack of interest was perhaps partly the result of Einstein's enormous influence and his strong preference for a simply connected space. Equally important, though, was the lack of any practical means to detect the topology of the universe experimentally. Cosmologists are most interested in questions that can be put to the test observationally. No practical tests were in sight, so cosmologists turned their attention elsewhere. Within a generation their ignorance of topological questions was so great

that most textbooks stated (incorrectly!) that a positively curved space must be a three-sphere, a flat space must be infinite Euclidean space, and a negatively curved space must be infinite hyperbolic space. All multiconnected manifolds had been forgotten.

Cosmologists rediscovered multiconnected manifolds during the last decade of the twentieth century, perhaps partly as a result of contact with mathematicians studying them intensely, but more likely because experimental tests were finally becoming practical. At the beginning of the twenty-first century, two distinct research programs are underway to test for a multiconnected universe: the method of Cosmic Crystallography (Chapter 21) looks for patterns in the arrangement of the galaxies, while the Circles in the Sky method (Chapter 22) examines microwave radiation remaining from the big bang in hopes of detecting the shape of our universe.

This Page Intentionally Left Blank

21

Cosmic Crystallography

In a multiconnected space we see multiple images of ourselves (Figures 7.1 and 7.2). So testing whether the real universe is multiconnected or not is easy, right? We just point our telescopes out into the night sky. If we see images of our Milky Way galaxy out there, then the universe is multiconnected. If we don't see images of the Milky Way, then either space is simply connected, or it's multiconnected but on too large a scale for us to observe it.

If only testing the topology of space were that easy! In reality there's one huge complicating factor:

the speed of light. The light we receive from distant galaxies is very, very old. If a distant galaxy is, say, a billion light years away from us, then the light we receive from it has taken roughly a billion years to reach us. We see the galaxy as it was roughly a billion years ago, not as it is today. Galaxies, like children, change a lot over the years, so even if we are seeing another image of our own Milky Way galaxy out there, we're seeing it as it was a billion years ago, so it's likely to look very different than it does now. Furthermore, if we're seeing different images of it in different parts of the sky, most likely we're seeing it from different angles. That is, we might see an image of it edge-on in one part of the sky, another image of it face-on in a different part of the sky, and so on.

The challenge of recognizing these diverse images as images of the same galaxy is comparable to the challenge of looking out into a crowd of a hundred billion people and trying to recognize a few dozen images of your mother. If her images were all identical the task would be straightforward. But think how difficult the task becomes if in one part of the crowd you see an image of your mother viewed face-on as a 20-year-old, while in another part of the crowd you see her as a 3-year-old viewed from the bottoms of her feet, while in yet another part of the crowd you see her as a 57-year-old viewed from the top of her head. It would be a challenge to recognize those three images as images of the same person. And keep in mind

that you have to distinguish her from the other hundred billion people in the crowd!

Exercise 21.1 Has the light from a galaxy a billion light years away from us taken *exactly* a billion years to reach us? Why or why not? □

Trying to recognize repeating images of the Milky Way is every bit as difficult as trying to recognize those images of your mother in the crowd. Fortunately Marc Lachièze-Rey, Roland Lehoucq, and Jean-Pierre Luminet have devised a way to test for repeating patterns without having to recognize individual galaxies. Their idea is quite simple: start with a catalog of galaxies (or other sources of light) and compute the distance between every pair of galaxies.

Exercise 21.2 In a catalog with only three galaxies (say A, B, and C) there would be only three distances to compute (AB, BC, and CA). In a catalog with four galaxies (A, B, C, and D) there would be six distances (AB, AC, AD, BC, BD, and CD). In a catalog with n galaxies, how many distances must be computed? □

In a simply connected universe (Figure 21.1) the computed distances aren't especially interesting. They obey a known statistical distribution (a so-called Poisson distribution) but are otherwise unremarkable. In a multiconnected universe, however, certain distances may occur more than once (Figure 21.2). Specifically, the distance between the two images of galaxy A is

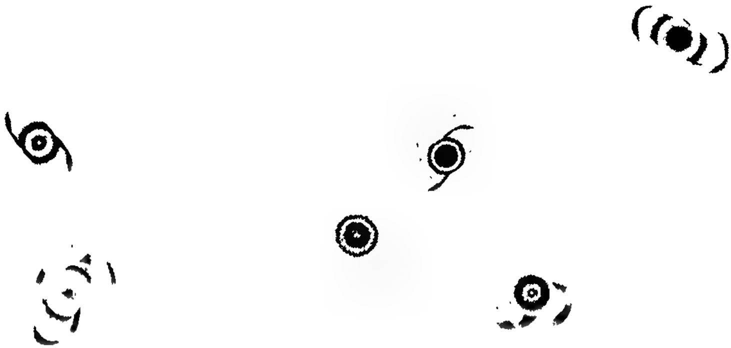


Figure 21.1 In a simply connected universe the arrangement of the galaxies is random.

exactly the same as the distance between the two images of galaxy B, which is in turn exactly the same as the distance between the two images of galaxy C. Experimentally, then, we may compute all possible distances between the galaxies in a galaxy catalog, and if we find that certain distances occur much more fre-

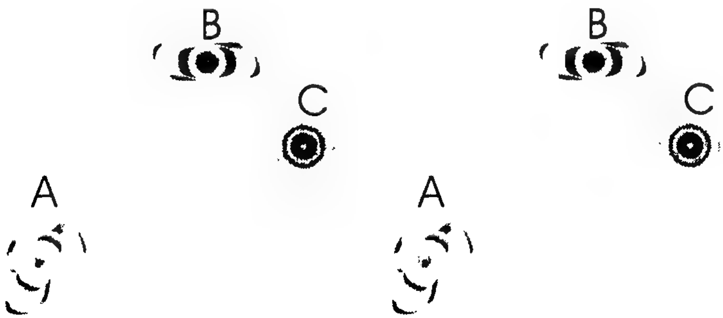


Figure 21.2 In a multiply connected universe images of galaxies repeat in a regular pattern.

quently than random chance would suggest, that is our clue that the universe is multiconnected. This method for detecting a multiconnected universe is called Cosmic Crystallography.

Problem Cosmic Crystallography doesn't work with galaxies. Galaxies are constantly in motion. For example, the Milky Way is moving through space at about 600 km/sec, and the other galaxies move at comparable speeds. So if we observe different images of the same galaxy in different parts of the sky, we see it not only at different points in its history, but also at slightly different points in space. Furthermore, because of observational uncertainties we don't know exactly where the observed images are in any case.

Resolution Don't work with individual galaxies. Work with superclusters of galaxies instead. Galaxies occur in clusters, and a supercluster is a cluster of clusters. We don't know the position of a supercluster any more accurately than we know the position of an individual galaxy. The point is that the same uncertainty in position, which is large compared to the size of an individual galaxy, is small compared to the size of a supercluster.

Working with superclusters has another practical advantage. There are roughly a hundred billion individual galaxies visible in the sky, but only a few hundred superclusters. Figure 21.3 shows a histogram of the expected separation distances between superclusters in a simply connected space. Figure 21.4 shows the same histogram for a three-torus. The second histogram is similar to the first, except for the spikes corresponding to the distances between two images of the same galaxy. The spikes reveal the multiconnectedness of the space. A more careful analysis of the data reveals the directions of the fundamental translations.

The downside of cosmic crystallography is that it doesn't work for all manifolds. For example, in a Klein bottle (Figure 21.5) the distance between two images of supercluster A needn't equal the distance between

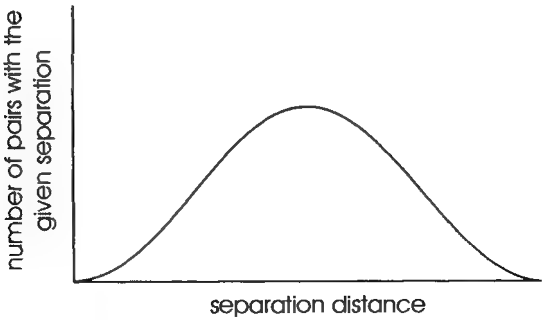


Figure 21.3 This histogram plots the frequency with which a given separation between superclusters occurs within the observable universe. Here we assume the universe is simply connected and all superclusters are distinct, like the galaxies in Figure 21.1.

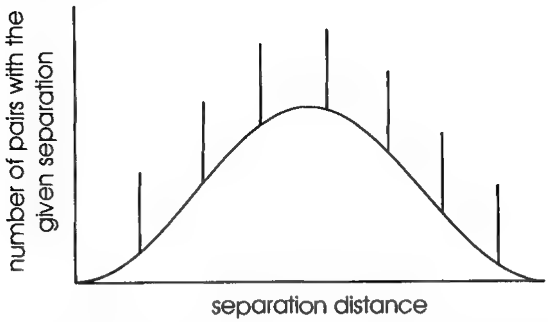


Figure 21.4 This histogram plots the frequency with which a given separation between superclusters occurs within the observable universe. Here we assume the universe is multiply connected and multiple images of the superclusters form a regular pattern, like the biplanes in Figure 7.3 and the galaxies in Figure 21.2.

two images of supercluster B. Because the crystallographic method fails to detect the reflected images, it cannot possibly recognize the Klein bottle. Instead it would detect only the pure translations (Figure 21.6) and falsely report the space to be a torus.

The situation in three dimensions is similar. For example, if the universe were a quarter turn manifold (Exercise 7.3), the crystallographic method would fail to detect the rotated images. It would detect only the pure translations and report the universe to be a three-torus with volume four times that of the original quarter turn space.

What about curved three-manifolds? In a hyperbolic three-manifold the distance between two images of supercluster A *never* exactly equals the distance between two images of supercluster B, so the crystallo-

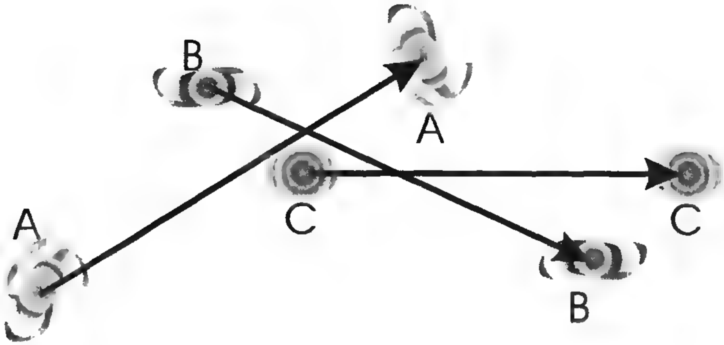


Figure 21.5 In Exercise 7.1 you found that the images in a Klein bottle form a repeating pattern in which every other image is a reflection of the one preceding it. In other words, neighboring images of the same object are related by glide reflections, not pure translations. The present figure shows part of the view in a Klein bottle containing galaxies A, B, and C. The second image of each galaxy is congruent to its first image under the action of a glide reflection. The distance from one image of galaxy C to the next is short, because galaxy C lies on or near the glide axis. The distance between the two images of galaxy B is longer, because galaxy B lies further from the glide axis. The distance between the images of galaxy A is slightly longer still, because galaxy A lies still farther from the glide axis.

graphic method fails entirely. In a spherical three-manifold the method detects some images but not others. Luckily the closest images are the ones most likely to be detectable, so if the real universe is a spherical manifold—and is small enough that we can see our nearest images—then the crystallographic method has an excellent chance of detecting its topology.

As of Autumn 2001, the crystallographic method has been applied only to very modest galaxy catalogs

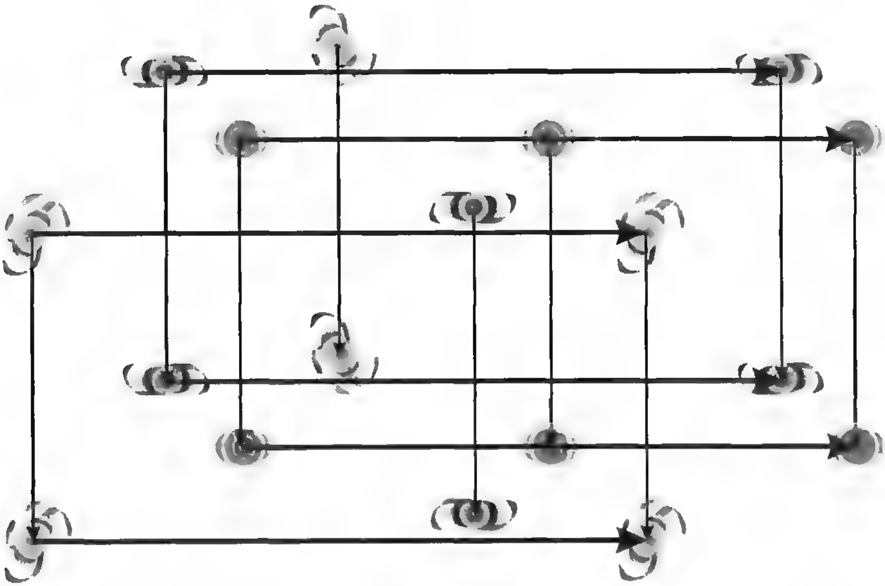


Figure 21.6 If you study your answer to Exercise 7.1 you'll notice that the view in a Klein bottle contains pure translations as well as glide reflections. Image pairs related by a pure translation are readily detectable by the crystallographic method. For example, in the present figure the heavy horizontal arrows mark six pairs of images sharing the same separation distance. The light vertical arrows mark nine different pairs of images that also share a common distance. The crystallographic method would detect the pure translations in a Klein bottle but miss the glide reflections, and would therefore erroneously report the space to be a torus with area twice that of the Klein bottle.

covering a limited volume of space, and no periodicity has been observed. More extensive catalogs will become available by about 2010, offering a better opportunity to detect the shape of space.

This Page Intentionally Left Blank

22

Circles in the Sky

During the first 300,000 years after the big bang, the universe was small and cramped, like the first two-holed torus near the top of Figure 19.2. The galaxies hadn't yet formed, and hydrogen and helium gas filled the whole universe. This was no ordinary gas, though. Because all the energy of the universe was packed into such a small volume, the temperature was extremely high, in excess of 3000 K. At such high temperatures, electrons get knocked loose from their atoms, and the gas takes the form of a so-called plasma, consisting of ions, electrons, and radiation. In other words, the

whole universe was filled with a hot glowing substance much like the outer layers of the modern sun. If you could somehow take a (heat resistant!) time machine back to that era, you'd find yourself in a blazing hot fog.

As the universe expanded, it cooled. About 300,000 years after the big bang, the universe had cooled enough that the hot plasma could finally condense to a gas. The universe became transparent, filled with warm but clear hydrogen and helium. What made the gas suddenly transparent? The technical reason is that photons of radiation scatter readily off charged particles in the plasma because they sense the particles' electrical charge (Figure 22.1), but the photons do not interact with the electrically neutral atoms in the gas (Figure 22.2).

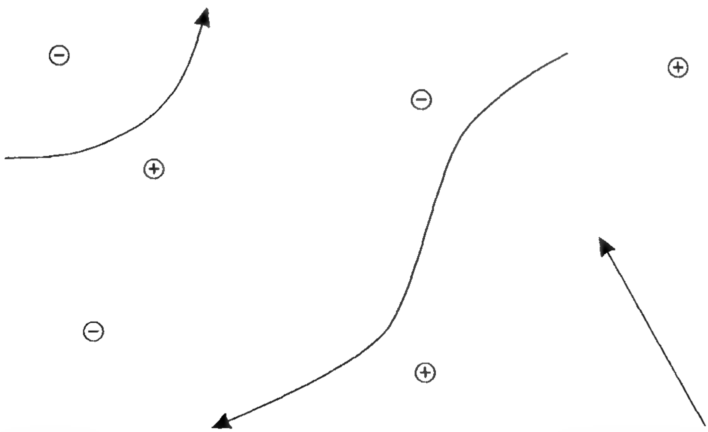


Figure 22.1 Photons in a plasma scatter off charged particles, so the plasma is opaque, like a hot glowing fog.

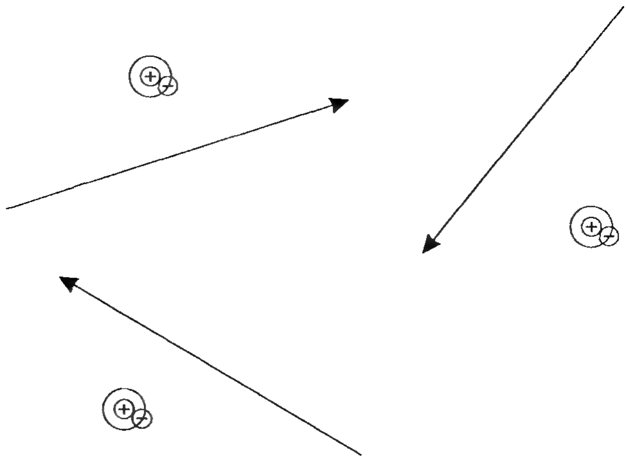


Figure 22.2 Photons do not interact with the atoms in a gas, so the gas is transparent.

The liberated radiation has been traveling more or less unimpeded across the universe ever since. Originally it had a temperature of about 3000 K, the same as the plasma at the time it began condensing to a gas. But the universe has expanded by a factor of about 1100 between then and now, and the 3000 K photons of infrared light have been stretched out to become microwaves at a chilly 2.7 K, or 2.7 degrees above absolute zero (Figure 22.3). This is the Cosmic Microwave Background (CMB) radiation. It presently fills the whole universe with a density of about 400 photons per cubic centimeter.

The CMB was first observed by Arno Penzias and Robert Wilson at Bell Labs in 1965. They weren't looking for it (they were instead trying to measure radio emissions from a supernova remnant), but when they

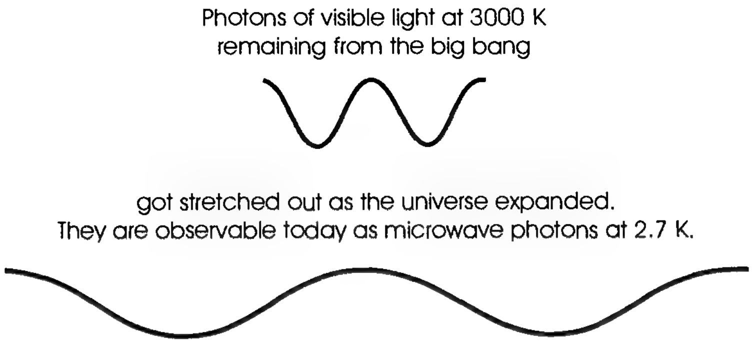


Figure 22.3

announced "excess antenna noise" the cosmologists immediately knew what it was, because the cosmologists were already constructing their own antenna.

The CMB has two striking properties. The first is its so-called thermal or blackbody spectrum, which was imprinted on it during the plasma era when radiation and matter were in thermal equilibrium. The curve in Figure 22.4 is a theoretical graph of a blackbody spectrum showing the intensity of the radiation as a function of the wavelength. In 1991 the COsmic Background Explorer (COBE) satellite measured the CMB at 43 wavelengths. You can see the data plotted in Figure 22.4, with nice tight error bars. These error bars are at the 400σ level (400 standard deviations from the mean). If you instead plotted the more traditional 2σ error bars, they would be completely invisible because they would all be smaller than the pixels used to draw the curve! This astonishingly precise blackbody spectrum—said to be the most precise

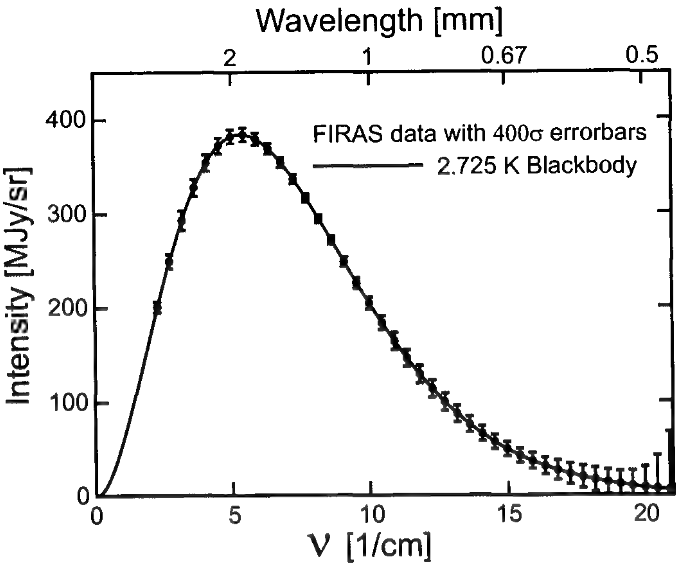


Figure 22.4 In 1991 the COBE satellite found the cosmic microwave background radiation to have a precise blackbody spectrum, exactly as the hot big bang theory had predicted. © Edward L. Wright, used with permission.

blackbody spectrum ever measured by humans—provides excellent confirmation that the hot big bang model (illustrated in Figure 19.2) is correct.

The second striking fact about the CMB is its uniformity: its temperature is constant across the sky to within a few parts in 10^5. This uniformity implies that the curvature of the observable universe is constant to within one part in 10^4. One must be careful, though, when drawing conclusions about the whole universe based on observations of the portion we see. It could well be that the whole universe has constant curvature—aesthetically this would be most satisfying.

However, it's equally plausible that the whole universe has varying curvature, and the portion we observe looks flat only because it's such a small percentage of the total volume. It's like measuring the surface of a frozen lake and determining that the Earth's surface is approximately flat; the only reason the lake appears flat is that it occupies such a tiny portion of such a huge sphere.

The uniformity of the CMB is impressive, but it's the small deviations from uniformity that may reveal the topology of space. To see how this is possible, let's take a closer look at the CMB fluctuations and what they reveal.

The CMB is homogeneous (it's distributed uniformly in space) and isotropic (it's heading in all directions), as illustrated in Figure 22.5. At first glance,

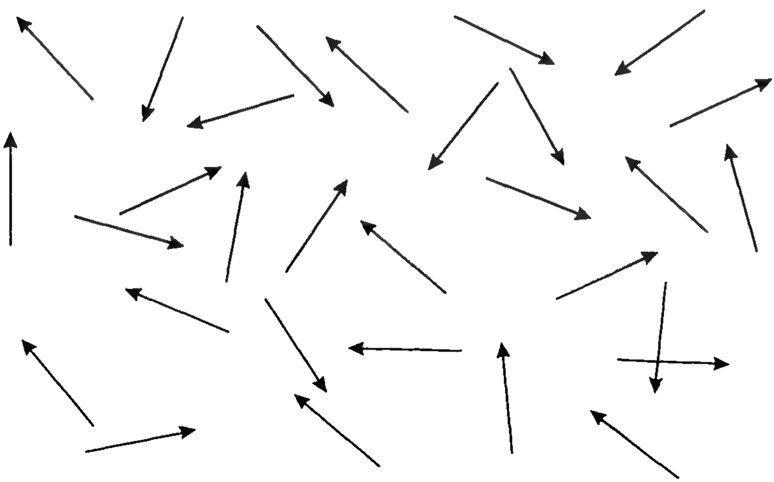


Figure 22.5 Cosmic microwave background photons fill space uniformly, traveling in all directions.

this doesn't seem like such a great source of information. But keep in mind that the CMB photons we observe with our detectors are the ones arriving *here* and *now*. These photons have all been traveling at the same speed (the speed of light) for the same length of time (since the primordial plasma condensed to a neutral gas). Therefore they have all traveled the same distance. This means that the CMB photons arriving on Earth today started their ten-billion-year voyage on the surface of a huge sphere, the so-called Last Scattering Surface (LSS) (Figure 22.6). This may seem anti-Copernican, to have humanity sitting grandly at the exact center of the LSS. Please keep in mind that the citizens of an extraterrestrial civilization in a dif-

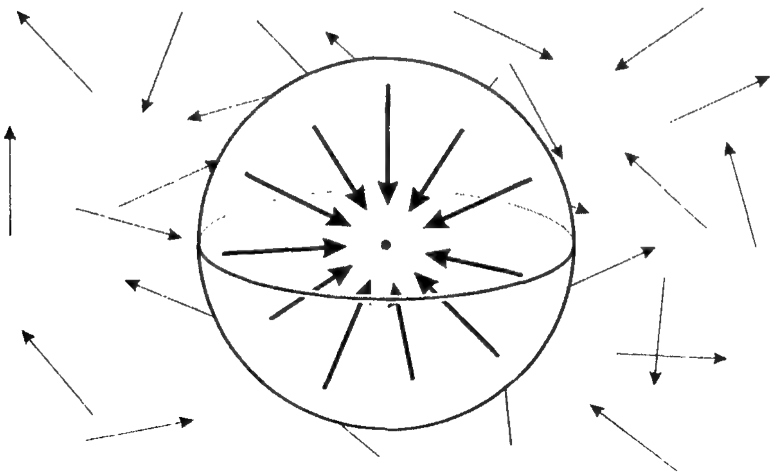


Figure 22.6 The CMB photons we observe are the ones arriving here and now. These particular photons began their 10-billion-year journey through space on the surface of a huge sphere called the Last Scattering Surface.

ferent part of the universe will observe their own CMB photons, which come to them from their own LSS, with their own civilization sitting grandly at the center. The LSS is defined relative to both the position of the observer and the time of the observation.

Exercise 22.1 Will our LSS be larger in the future? Smaller in the future? The same size for all time? □

When we observe the CMB we are literally looking back in time and seeing the primordial plasma. At that time, roughly 300,000 years after the Big Bang, the universe was a very homogeneous place, but it wasn't perfectly homogeneous. There were small density fluctuations to the order of one part in 10^5. Photons coming to us from denser regions in the plasma do a little extra work against gravity, and so they cool a little more than average during their ten-billion-year trip. Conversely, photons coming to us from less dense regions do less work against gravity, and they arrive a little warmer than average. Thus temperature fluctuations in the CMB reveal density variations in the primordial plasma.

How do the fluctuations tell us the shape of space? Figure 22.7 shows the LSS in a three-torus universe. There is only one LSS, but we see multiple images of it, just as we saw multiple images of the biplanes in Figures 7.1–7.3 and multiple images of Einstein in Figure 7.4. In Figure 22.7 the universe is slightly larger than the LSS, and so we learn nothing about its topology.

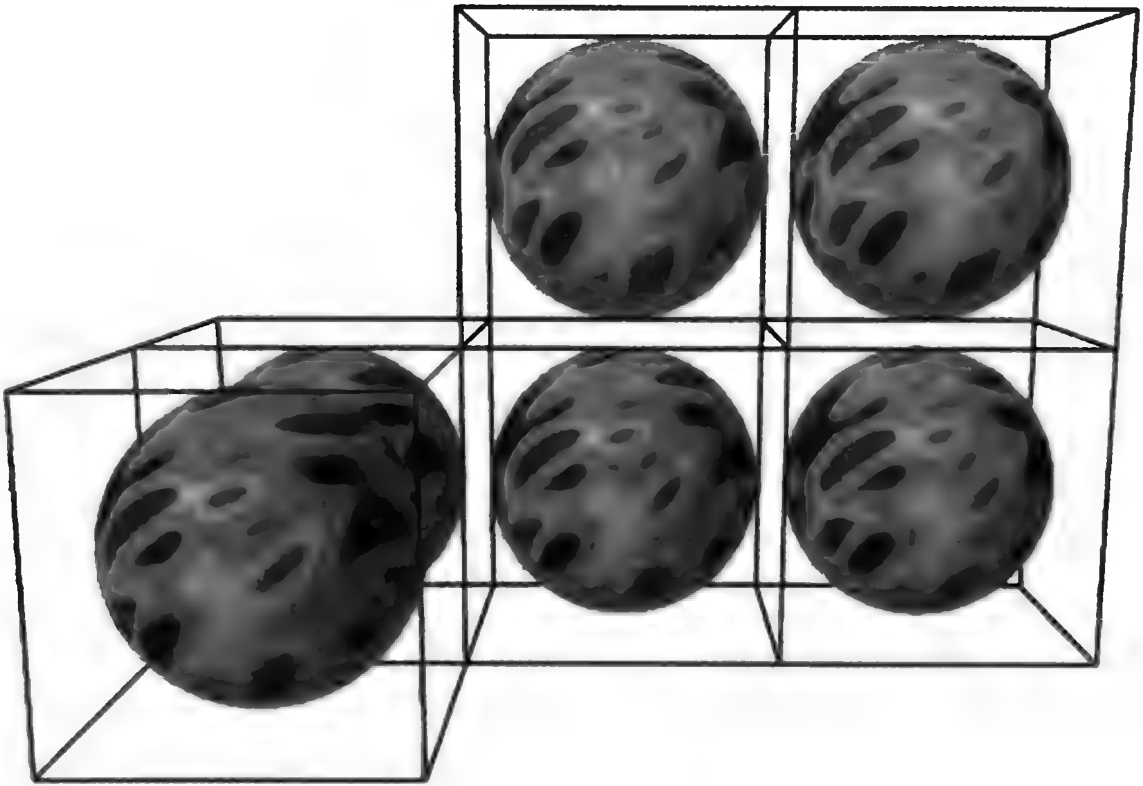


Figure 22.7 If the universe is too big, then the last scattering surface does not intersect itself.

The situation in Figure 22.8 is more fortunate. Here the three-torus universe is slightly smaller than the LSS. The LSS wraps around the universe and intersects itself. Each self-intersection is a circle. We humans, sitting at the center of the LSS, can look, say, to the west and see one of the circles of self-intersection. We can also look to the east and see the *same* circle of self-intersection. That is, the same circle of points in space appears once in the western sky and once in the eastern sky.

Figure 22.9 shows the microwave sky as seen by an observer at the center of the LSS in Figure 22.8.

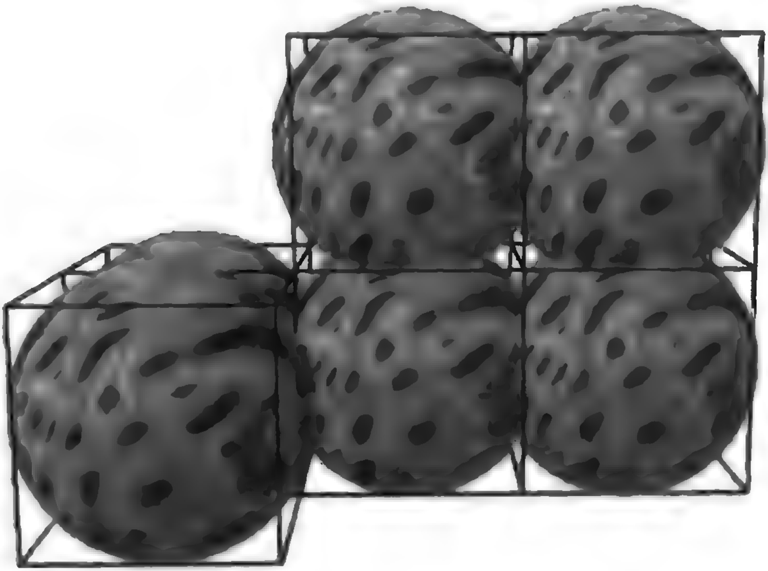


Figure 22.8 If the universe is smaller than the last scattering surface, then the last scattering surface overlaps itself.

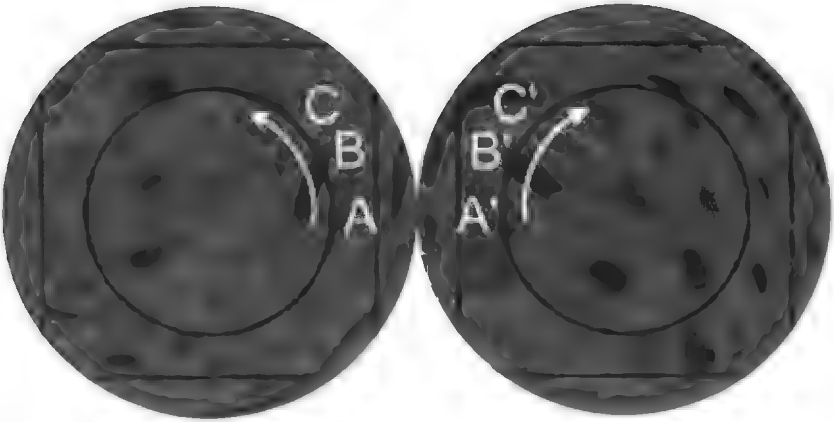


Figure 22.9 As you trace your finger around the circle in the western sky (left) you'll pass the same temperatures as you would at the corresponding points on the circle in the eastern sky (right). The overall temperature patterns in the two hemispheres are very different—the temperatures match only along the circles.

The sky is a sphere and it's awkward to draw a sphere on the flat page of a book, so Figure 22.9 splits the sky into western and eastern hemispheres. The circles of self-intersection are marked. Because the circle centered in the western sky represents the same points in space as the circle centered in the eastern sky, we expect corresponding points to have equal temperatures.

Don't be misled by the fact that the circle in the western sky runs counterclockwise while the circle in the eastern sky runs clockwise. If you imagine the two hemispheres to be joined with a hinge, and you imagine folding them closed to restore a spherical sky with

you sitting at the center, then you will see that the two circles are directly opposite each other in three-dimensional space and run in the same direction.

Exercise 22.2 (a) Trace around the circle centered in the western sky of Figure 22.8 using your left index finger, starting at point A and going counterclockwise, while simultaneously tracing around the circle in the eastern sky with your right index finger, starting at point A′ and going clockwise. Do your fingers pass over equal temperatures at corresponding points?
(b) Locate the circles centered in the northern and southern skies. Do they also have equal temperatures at corresponding points? □

Exercise 22.3 (a) If the universe of Figure 22.8 were a quarter turn manifold instead of a three-torus, how would that affect the matching circles?
(b) If it were $K^2 \times S^1$, how would that affect the matching circles? □

Exercise 22.4 How would the matching circles appear in the Poincaré dodecahedral space? How about in the Seifert–Weber space? By studying the matching circles alone, could you decide whether you were living in a Poincaré universe or a Seifert–Weber universe? □

Neil Cornish, David Spergel, and Glenn Starkman were the first people to realize that in a sufficiently small multiconnected universe the last scat-

tering surface intersects itself, and the circles of intersection reveal the shape of the space.

At 3:46 PM on June 30, 2001, a Delta II rocket lifted NASA's Microwave Anisotropy Probe (MAP) (Figure 22.10) into space. Its acronym "MAP" is most appropriate, because MAP will spend two years making a very accurate map of the CMB. A few years later the European Space Agency will launch the Planck satellite, with a similar purpose and even greater ac-

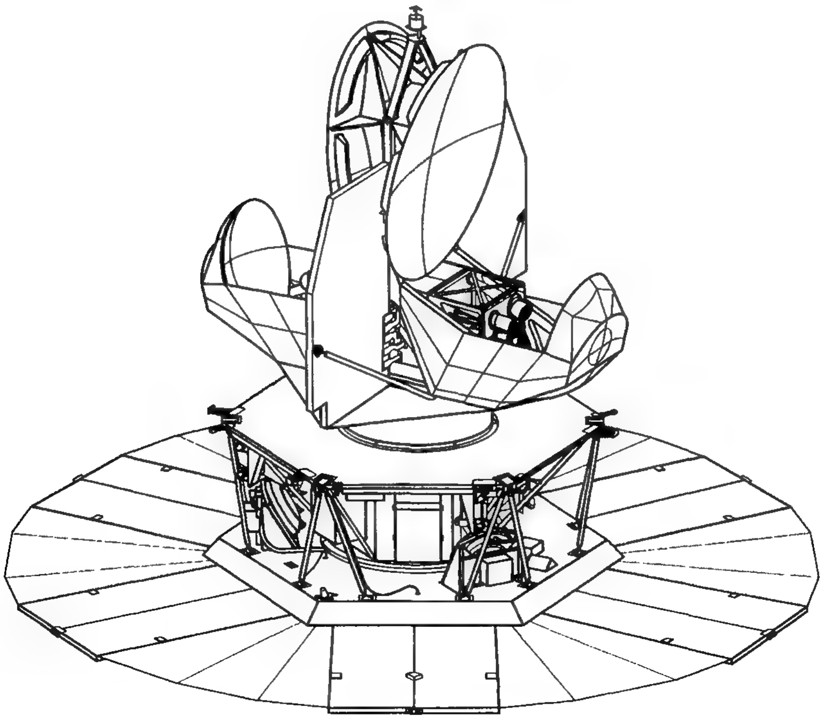


Figure 22.10 The Microwave Anisotropy Probe (MAP) will make a full-sky map of the cosmic microwave background radiation. Courtesy of NASA MAP Science Team.

curacy. The main purpose of these sky maps is not to determine the shape of space, but to deduce accurate values for the rate of expansion, the curvature of space, and the amount of vacuum energy, and to study how the galaxies formed. Luckily MAP's map of the microwave sky will serve equally well to search for topology. Researchers will carry out a massive computerized search, scanning the microwave sky for pairs of circles with nearly identical temperature distributions. If the universe is larger than the LSS, then of course no matching circles will be found. But if the universe is small enough that the LSS overlaps itself, and the noise in the MAP data or eventually the Planck data isn't too severe, then we will find matching circles, and from them deduce the shape of space.

Appendix A

Answers

Chapter 2

2.2 Player X can win by taking the middle square of the bottom row. Player O wins by taking the bottom left square.

2.3 Positions a, b, c, e, and g are all equivalent, as are positions d, f, and h. Position c is obtained from a by moving everything down one notch. Position g is obtained from a by moving everything to the left one notch. Position b is obtained from a by moving everything down a notch *and* to the left a notch. Position e comes from rotating a a quarter turn. Position f is

obtained from d by rotating it a half turn (or by moving it down and to the right), and h is obtained from f by moving everything one notch to the right.

2.4 The first player has, in effect, only one possible move: any move she makes can be shifted upward and/or to the right until it appears in the center of the board. Her opponent has two possible responses: either directly next to her or on a diagonal. With optimal play the first player can always win.

2.5 The white knight threatens a bishop, a king, and two knights, while a queen, a rook, a bishop, and the knights threaten it.

2.6 Every black piece is threatened by both the white knight and the white queen.

2.8 The lower left-hand corner. (Imagine him going first up one square and then to the right one square.)

2.9 They cannot, for the same reason as in conventional chess: a bishop threatens only pieces on the same color square, while a knight threatens only pieces on an oppositely-colored square.

2.10 When you look through a wall you see what appears to be another copy of the room. You see yourself from behind. When you look through the floor you see the top of your head, and when you look through the ceiling you see the bottoms of your feet.

2.12 Throw the ball and turn around quick. If you're fast enough you can even get in a little batting practice.

Chapter 3

3.1 Surfaces *a* and *c* have the same topology. So do *b* and *d*, and *e*, *f*, *g*, and *h*. Surface *h* may look like it has four holes, but if you flatten it out you'll see that it has only three. For *g* you need to deform the surface to uncross the cross pieces.

3.2 He discovered a topological property. The property remains even if Flatland is distorted.

3.3 Surfaces *a*, *c*, *d*, and *f* have the same extrinsic topology, as do surfaces *b* and *e*. The hardest part is seeing that *c* and *f* are the same as *a*: shrink the connecting piece in *f* to make it look like *c*, and then pull the inner loop out from the outer loop to make *c* look like *a*. Good luck!

3.4 Reglue the band with a half-twist instead of a full twist. (You get a "Mobius strip".)

3.5 You can roll the paper into a cone, but you can't wrap it smoothly onto a basketball. Try it! If you don't have a basketball, use somebody's head. The paper, the cylinder and the cone have the same intrinsic geometry, but the intrinsic geometry of the basketball is different.

3.6 Imagine the universe as a room with opposite walls glued. Home is next to one wall, your friends

are in the middle, and you do your exploration over by the opposite wall. After you're done exploring you simply pass through the wall and you are home.

3.7 Discovery #1 is local, while discoveries #2 and #3 are global.

3.8 A one-dimensional manifold is a space with the local topology of a line. A circle is an example of a one-dimensional manifold. (In fact, a circle is the *only* "closed" one-dimensional manifold.)

3.9 The cylinder and the plane have the same intrinsic local geometry (but different extrinsic local geometries). They have different global topologies, but they have the same local topology as do all surfaces. Only discovery #2 might distinguish a cylinder from a plane, and even then Flatlanders on a cylinder would have to travel in exactly the right direction.

3.10 They have the same local topology (as do all three-manifolds), the same local geometry (namely that of ordinary Euclidean space), and the same global topology (one can easily be deformed to the exact shape and proportions of the other). However, the fact that their overall dimensions are different means that they have different global geometries.

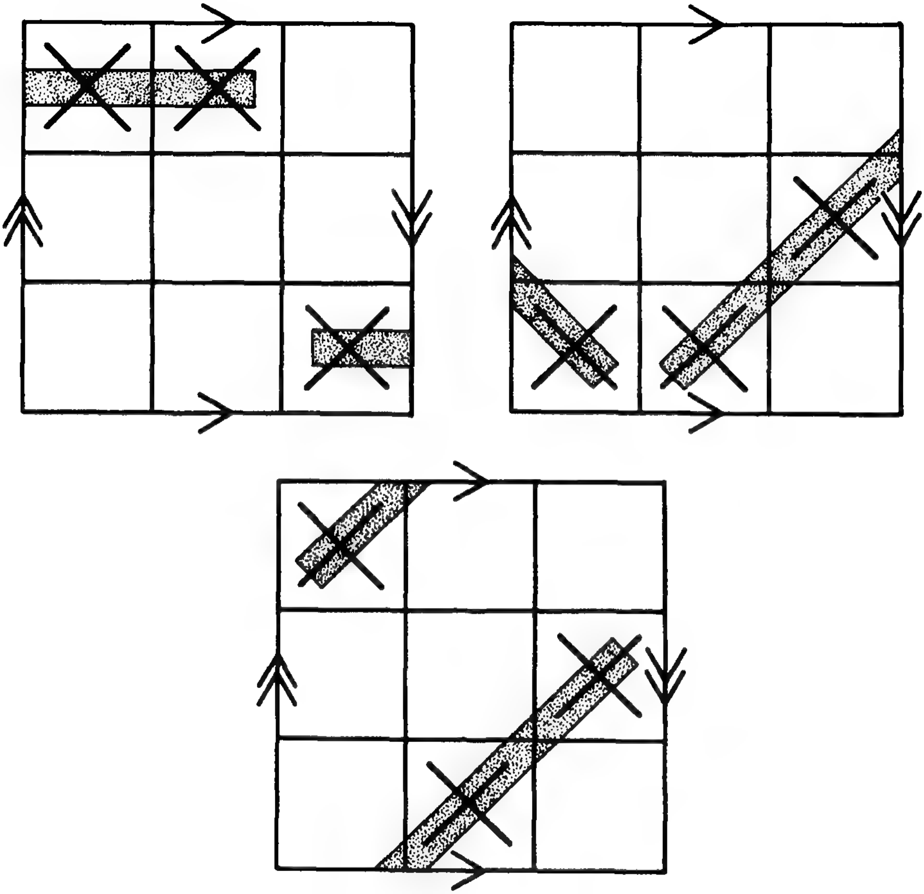
3.11 Yes. At any point it has the local geometry of ordinary Euclidean space.

3.12 1, 3, 4, 7 and 9 are closed; 2, 5, 6 and 8 are open.

3.13 He bumps into himself.

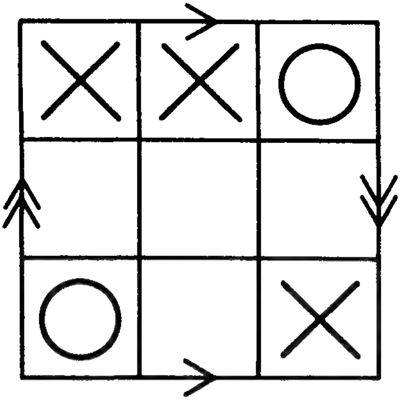
Chapter 4

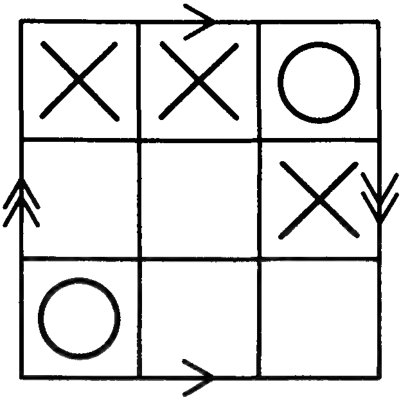
4.2 They all do!

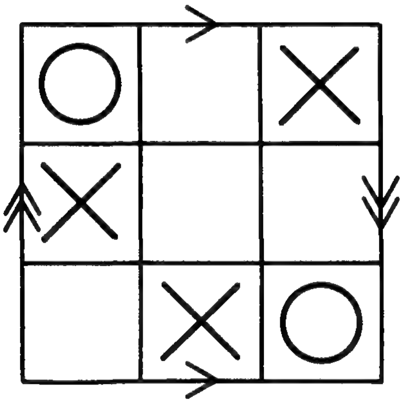


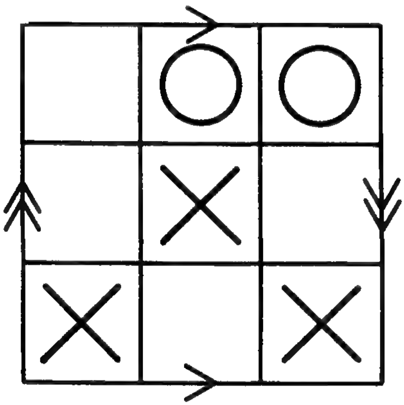
4.3 The white knight threatens a bishop, a knight and a rook. The same bishop, knight and rook threaten it.

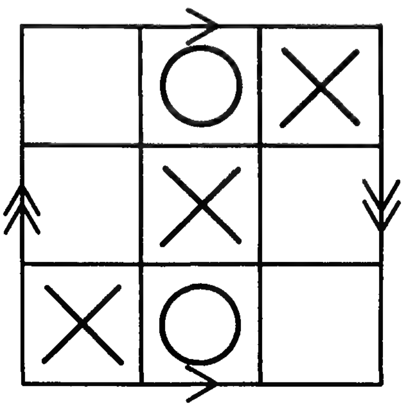
4.4 On all boards X can win immediately. See page 314. (Incidentally, on each of the first four boards X will lose if he tries to block O, because O has two separate ways to win.)











4.7 The bishop returns from the *upper* left hand corner.

4.8 Each square on the right side of the chessboard is adjacent to a square *of the same color* on the left side of the board. So when the bishop passes from one side of the board to the other it returns on a different color square. This phenomenon no longer occurs on the Klein bottle chessboard of Figure 4.8, because there every square is adjacent only to squares of the opposite color. A knight and a rook can still simultaneously threaten each other on this new board: place the rook in the row just above the center, and place the knight one space over and two spaces down.

4.9 When you look through the back wall you see a mirror-reversed copy of the room. When you look through the other walls you see normal copies of the room, just like in the three-torus.

4.11 The projective plane is nonorientable: A Flatlander crossing the rim comes back with his left and right sides interchanged.

4.12 He's furthest from home when he's crossing the rim. After that he's getting closer again.

4.13 Two fire stations should be positioned "90° apart." For example, they could be 90° apart on the rim, or one could be on the rim with the other at the south pole. Three fire stations should be positioned so that any two are 90° apart. One way to do this is to

have two of them 90° apart on the rim with the third at the south pole.

4.14 To decide whether he is on a projective plane or a sphere, a Flatlander can walk in a straight line until he gets back to his starting point: on a projective plane he comes back as his mirror image, but on a sphere he doesn't. This trick won't work to tell a projective plane from a Klein bottle; instead the second Flatlander can measure the angles of a triangle and see whether they add up to 180° or not.

4.15 Sphere: curved and orientable. Torus: flat and orientable. Klein bottle: flat and nonorientable. Projective plane: curved and nonorientable.

4.16 Projective three-space is orientable. If you cross the "seam" you come back rotated 180°, but not mirror-reversed. For more explanation, see Figure 14.8 and the corresponding paragraph in the text of Chapter 14.

4.17 Orientability is a global property because it says something about a manifold as a whole. It is a topological property because deforming a manifold does not affect it.

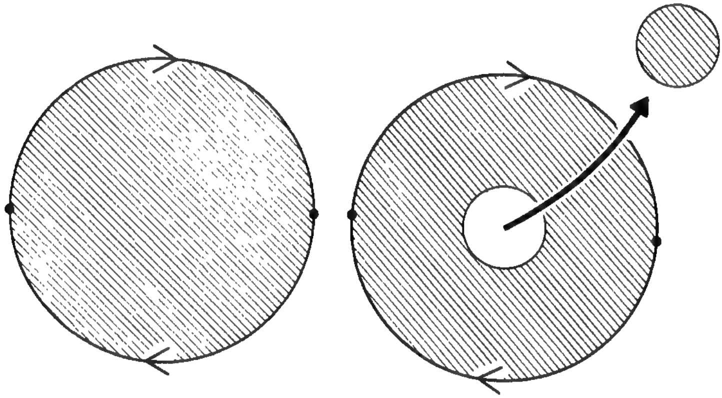
Chapter 5

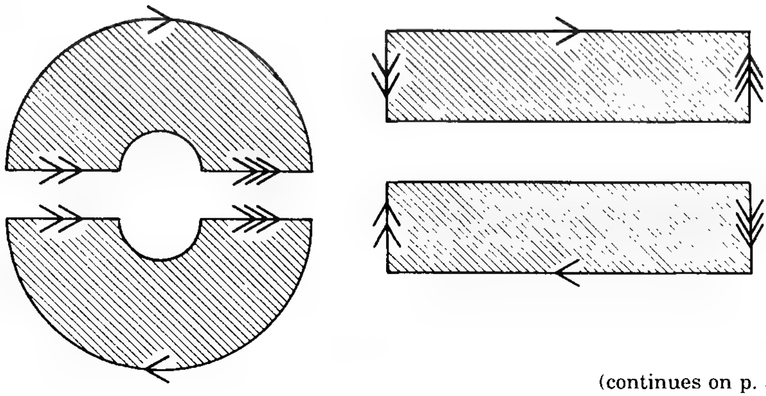
5.1 The connected sum of a two-holed doughnut surface and a one-holed doughnut surface is a three-holed doughnut surface (you can make a drawing just like Figure 5.4). Similarly, a six-holed doughnut surface

and an eleven-holed doughnut surface form a seventeen-holed doughnut surface.

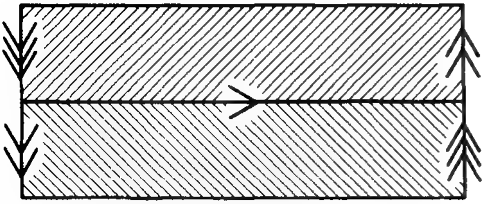
5.2 The connected sum of any surface and a sphere is just the same surface you started with. The reason is that a sphere with a disk cut out is topologically a disk, so when you form the connected sum you are removing one disk from your surface and replacing it with a different one.

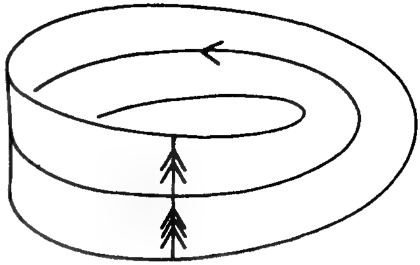
5.3 A projective plane with a disk cut out is a Möbius strip.





(continues on p. 318)





5.4 The connected sum of two projective planes is a Klein bottle. You start with two projective planes, cut a disk out of each to get two Möbius strips (as per Exercise 5.3), and then glue the Möbius strips' edges together to get a Klein bottle (as per Figure 5.6).

5.5 Flatland was the connected sum of a torus and a projective plane. The safe region was a torus with a disk cut out, and the reversing region was a Möbius strip, which is a projective plane with a disk cut out.

5.6 $(T^2 \# T^2) \# S^2 = T^2 \# T^2$, $K^2 \# S^2 = K^2$ and $P^2 \# S^2 = P^2$. (In general for any surface X, $X \# S^2 = X$.)

5.7 (a) $K^2 \# P^2 = P^2 \# P^2 \# P^2$, (b) $K^2 \# T^2 = P^2 \# P^2 \# T^2 = T^2 \# P^2 \# P^2$, and (c) $K^2 \# K^2 = P^2 \# P^2 \# P^2 \# P^2$.

5.9 Exercise 5.8 says that $T^2 \# P^2 = K^2 \# P^2$, and Exercise 5.4 implies that $K^2 \# P^2 = P^2 \# P^2 \# P^2$.

5.10 A surface written as a connected sum of both tori and projective planes can always be rewritten as a connected sum of projective planes only. Without

changing the surface's global topology you can convert the tori to Klein bottles (by Exercise 5.8) and then convert the Klein bottles to projective planes (by Exercise 5.4). Note, though, that you cannot convert a surface consisting of tori only into a surface consisting of projective planes only, because Exercise 5.8 does not apply when no projective planes are present. (Besides, you know a connected sum of tori cannot be the same as a connected sum of projective planes because one is orientable while the other is not.) $T^2 \# P^2 = K^2 \# P^2 = P^2 \# P^2 \# P^2$, $T^2 \# K^2 = T^2 \# P^2 \# P^2 = K^2 \# P^2 \# P^2 = P^2 \# P^2 \# P^2 \# P^2$, $P^2 \# S^2 = P^2$ and $S^2 \# S^2 = S^2$. The sphere and the connected sums of tori are all orientable, while the connected sums of projective planes are all nonorientable.

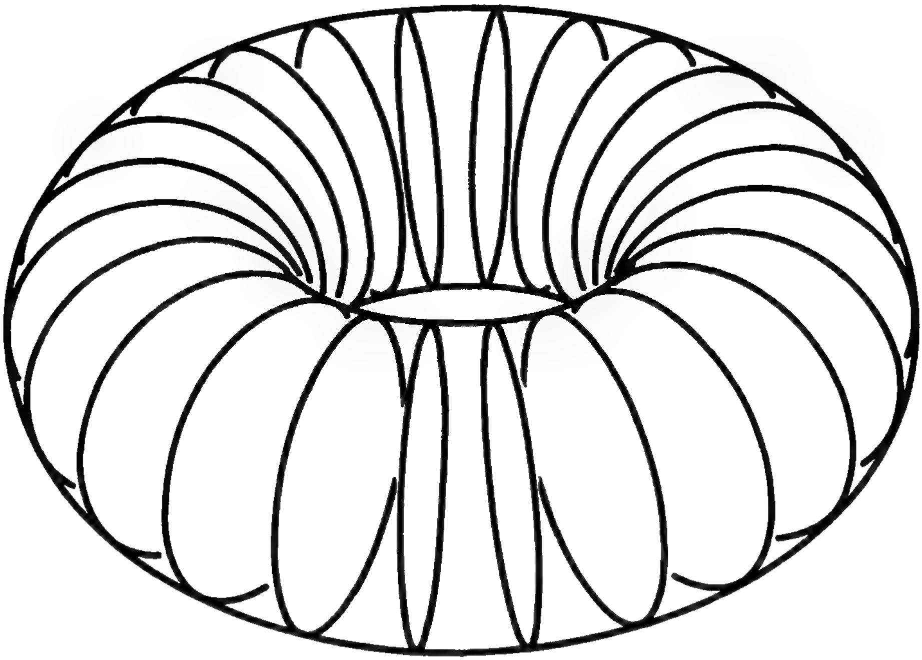
5.11 $T^2 \# S^2 = T^2$, $K^2 = P^2 \# P^2$, $S^2 \# S^2 \# S^2 = S^2 \# S^2$, $P^2 \# T^2 = K^2 \# P^2$ and $K^2 \# T^2 \# P^2 = P^2 \# P^2 \# P^2 \# P^2 \# K^2$.

Chapter 6

6.1 1. Square, 2. Plane, 3. Infinite cylinder, 4. Infinite strip.

6.2 No: it's a circle of intervals but it's not an interval of circles. (Chapter 17 will address this issue in more detail.)

6.3 $D^2 \times S^1$ is topologically a solid doughnut. The picture on p. 320 shows that it's a circle of disks; it's also a disk of circles, but that's a little harder to draw.



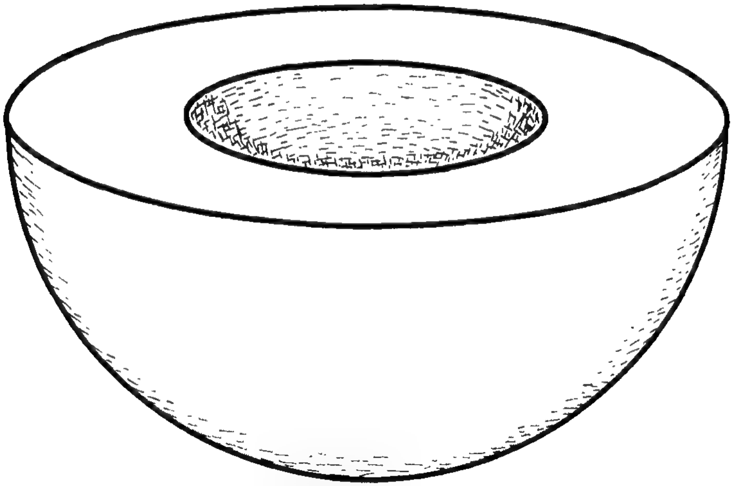
6.4 In the first example the circles are not all the same size, in the last example the intervals are not all the same size, and in the middle example the circles are not perpendicular to the intervals.

6.5 The geometrical product is a perfect square. *Any* deformed square will be only a topological product.

6.6 It's a Klein bottle cross a circle ($K^2 \times S^1$). (Each horizontal layer in Figure 6.5 becomes a Klein bottle when the cubes sides are glued.) It's a geometrical product because the horizontal Klein bottles are all the same size, the vertical circles are all the same size,

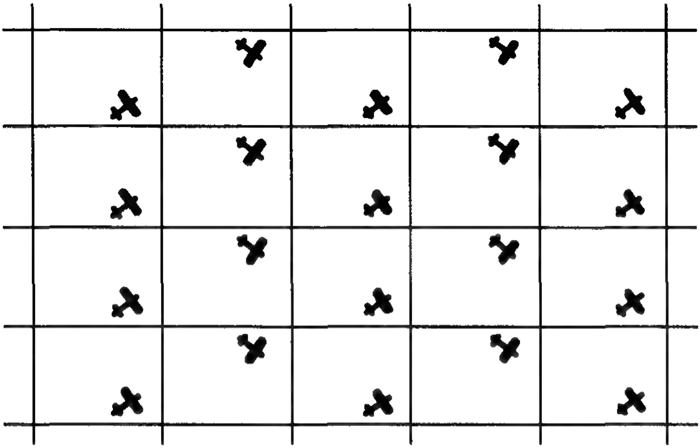
and the Klein bottles and the circles are perpendicular to each other.

6.7 $P^2 \times S^1$. You can visualize $P^2 \times S^1$ as a thickened hemisphere, as shown below. The inside surface is glued to the outside surface in the obvious way, and points on the "rim surface" are glued so that each hemispherical layer becomes a projective plane. This manifold is locally identical to $S^2 \times S^1$, but its global topology is different. Can you find an orientation reversing path?

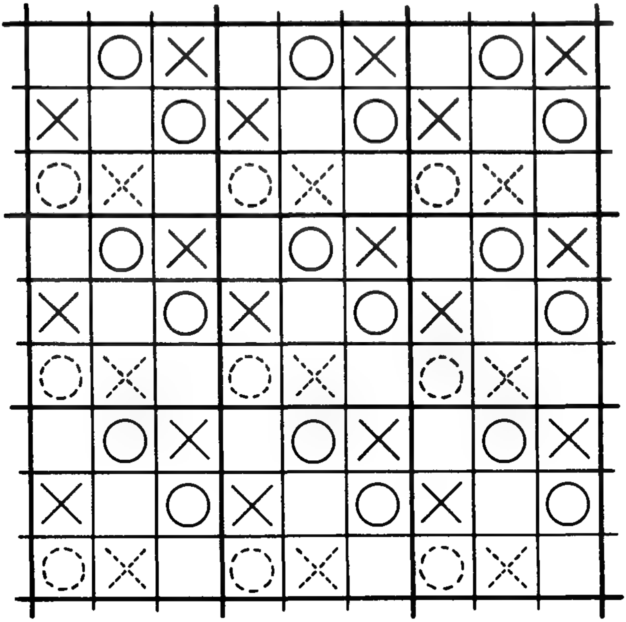


Chapter 7

7.1



7.2



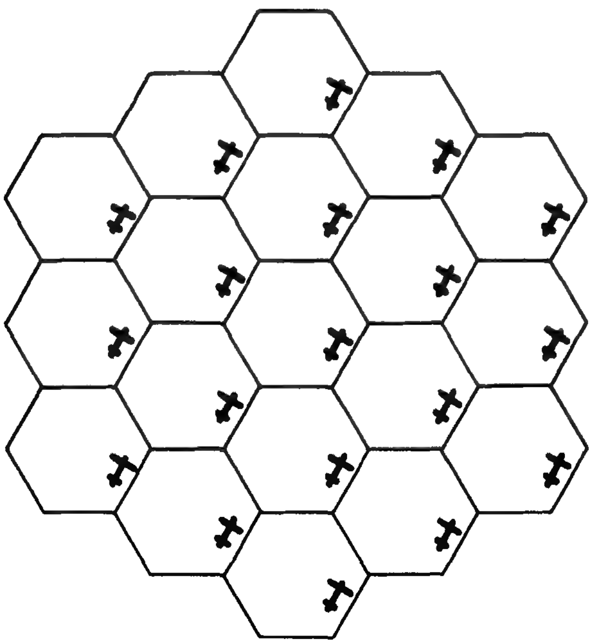
7.3 In the quarter turn manifold you see what appears to be an infinite lattice of cubes: all the cubes on a given level are colored identically, all the cubes on the next higher level appear to be rotated a quarter turn, the cubes on the level after that are rotated an additional quarter turn, etc. The views in the half turn and three-quarters turn manifolds are analogous. The one-quarter turn and three-quarters turn manifolds are mirror images of each other. Surprisingly, you can't tell one from the other intrinsically! The reason is that you have no standard by which to judge clockwise and counterclockwise; looking at the hands on your watch will do no good because you don't know whether you showed up in the manifold as your "normal self" or your "mirror image." Philosophically inclined readers may wish to ponder this matter some more: the heart of the problem is that in everyday speech we use terms like clockwise and counterclockwise, and right and left, in an absolute way, but really they are relative terms and can be used only to compare one object to another object at the same location.

7.4 Figure 7.10: The front is glued to the back with a side-to-side flip, the top is glued to the bottom with a side-to-side flip, but the left side is glued to the right side normally. Figure 7.11: The front is glued to the back with a side-to-side flip, the top is glued to the bottom with a front-to-back flip, and the left side is glued to the right side normally. Figure 7.12: The front is glued to the back with a side-to-side flip, the top is

glued to the bottom with a half turn, and the left side is glued to the right side normally. (You are correct if you said that the top is glued to the bottom in 7.12 with both a side-to-side and a front-to-back flip, because this is the same as a half turn.)

7.5 The manifolds in Exercise 7.3 are all orientable; if you go on a trip in one of them you might come back rotated but you won't come back mirror reversed. The manifolds in Exercise 7.4 are all nonorientable; in each one at least one pair of faces (e.g. the front and back) are glued with a flip.

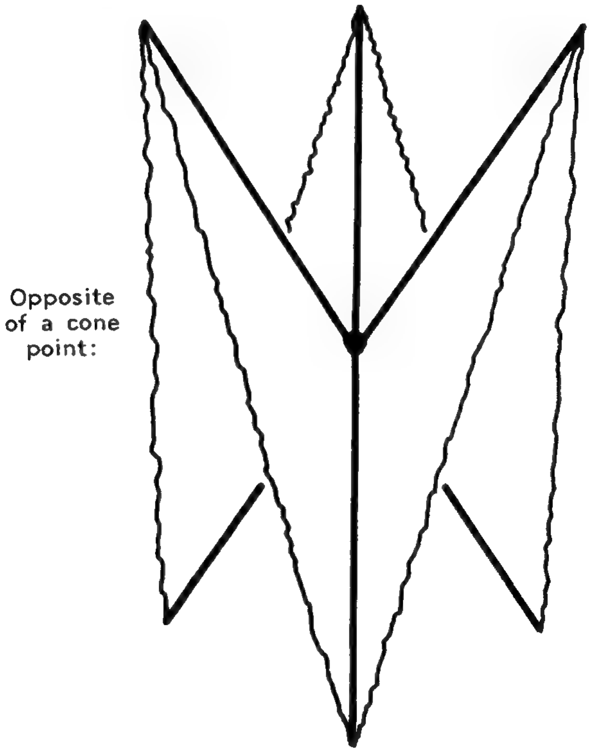
7.6



If a second biplane is flying around in the hexagonal

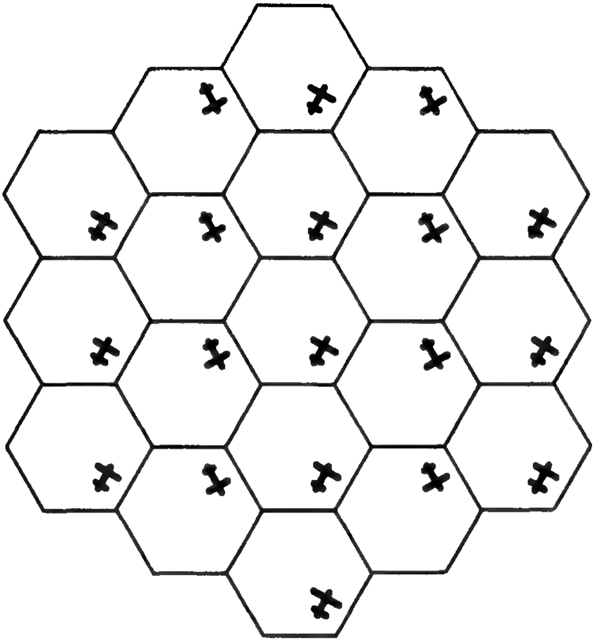
torus, then of course it too will appear once in each hexagonal cell.

7.7 In the first surface the hexagon's corners meet in three groups of only two corners each, so the surface has three cone points. In the second surface the hexagon's corners fit perfectly in two groups of three corners each; there are no cone points. In the third surface the hexagon's six corners meet at a single point! You get the opposite of a cone point: there is too much angle surrounding the vertex instead of too little.



7.8 None of them.

7.9 The surface is a hexagonal Klein bottle.



7.10 The hexagonal torus and the ordinary flat torus have the same local topology (as do all surfaces), the same local geometry (both are flat) and the same global topology (both can be deformed to a doughnut shape). On the other hand they have different global geometrical properties, as evidenced by the pattern of the images in Exercise 7.6.

7.11 Only the second surface counts as flat. The first is disqualified by its cone points, and the third by its opposite-of-a-cone-point.

7.12 The hexagonal three-torus and the usual three-torus have the same local geometry; both are flat, that is, they have the same local geometry as ordinary Euclidean space. They both have the same global topology because each is a torus cross a circle (in one case it's a hexagonal torus cross a circle and in the other it's an ordinary flat torus cross a circle, but those are geometrical differences, not topological ones). Their global geometries are different for the reason just mentioned. When you look around in a hexagonal three-torus you see copies of yourself arranged in layers; each layer forms a hexagonal lattice.

7.13 You'd see copies of yourself arranged in layers, and each layer would still be a hexagonal lattice just like in the hexagonal three-torus. Only now the copies of you in each layer are rotated a one-third (or a one-sixth) turn relative to the copies in the layer below.

Chapter 8

8.1 (1) Torus, orientable and two-sided. (2) Klein bottle, nonorientable and two-sided. (3) Torus, orientable and one-sided. (4) Klein bottle, nonorientable and one-sided.

8.2 Sideness is an extrinsic property of a surface because it has to do with how the surface is embedded in a three-manifold (for a specific example, note that the tori in the first and third drawings of Figure 8.2 are intrinsically identical, even though one is two-sided and the other is one-sided). Orientability is an

intrinsic property because Flatlanders living in a surface can tell whether it's orientable or not. All Klein bottles are nonorientable, but some are one-sided (e.g. the one in the last drawing of Figure 8.2) while others are two-sided (e.g. the one in the second drawing of Figure 8.2).

8.3 Use the center strip (running front to back) of the two-sided Klein bottle in the second drawing of Figure 8.2.

Chapter 9

9.1 First triangle: area = $\pi/2$, angle-sum = $3\pi/2$. Second triangle: area = $\pi/6$, angle-sum = $7\pi/6$. Third triangle: area = $2\pi/3$, angle-sum = $5\pi/3$. Fourth triangle: area = π, angle-sum = 2π. Fifth triangle: area = $\pi/2$, angle-sum = $3\pi/2$. Sixth triangle: area = 2π, angle-sum = 3π. The formula relating the area to the angle-sum appears later in the text.

9.2 The sum of the angles of the first triangle is $\pi/2 + \pi/3 + \pi/4 = 13\pi/12$, so its area must be $13\pi/12 - \pi = \pi/12$. For the second triangle you must convert each degree measurement to radians by multiplying by $\pi/180$. The area works out to be $A = (1.065 + 1.082 + 1.100) - \pi = 3.247 - 3.142 = 0.105$. (Computational note: when the angles are given in degrees, it is often more efficient to subtract 180° from the angle sum before converting to radians. Thus in this problem you would first compute $(61° + 62° + 63°) - 180° = 6°$, and then convert 6° to radians to get the answer, $6° \times (\pi/180) = 0.105$.)

9.3 *Direct method*: Modify the book's derivation of the formula $A = (\alpha + \beta + \gamma) - \pi$ to apply on a sphere of radius two, radius three or radius r. Start by finding the correct formula for the area of a double lune, and work from there. Your final result, for a triangle on a sphere of radius r, should be $A = r^2[(\alpha + \beta + \gamma) - \pi]$. *Fancy method*: Consider a triangle on a sphere or radius one. Now enlarge the sphere by a factor of two, and let the triangle get enlarged along with it. The angles of the triangle do not change, but its area increases by a factor of four (because both its height and its width have doubled). In general, when you increase the radius of the sphere by a factor of r, the area of the triangle, just like the area of the whole sphere, increases by a factor of r^2. In other words, a triangle on a sphere of radius r has r^2 times the area of a similar triangle on a sphere of radius one. The formula for its area is therefore $A = r^2[(\alpha + \beta + \gamma) - \pi]$.

9.4 The field's area is $A = r^2[(\alpha + \beta + \gamma) - \pi] = (1000\text{ m})^2[(0.76138 + 1.48567 + 0.89483) - 3.14159] = (1{,}000{,}000\text{ m}^2)[0.00029] = 290\text{ m}^2$. You start off with five significant figures, but after the subtraction you are left with only two! Thus you know the correct answer only to within a few percent. What if the original data had had only three digits of accuracy?

9.5 Plug into $A = r^2[(\alpha + \beta + \gamma) - \pi]$ to find the sphere's radius, which works out to be r = 10,000 meters. The sphere's area is therefore $A = 4\pi r^2 = 1.2 \times 10^9$ square meters.

9.6 (1) $\pi + (2 \times 10^{-5})$ radians, (2) $\pi + (.01)$ radians, (3) $\pi + \frac{1}{2}$ radians.

9.7 A projective plane has the same local geometry as a sphere, and the formulas for spherical triangles apply to it too. Flatlanders on a projective plane couldn't tell locally that they weren't on a sphere, but they could tell globally because a projective plane is nonorientable, and long-distance travellers can come back mirror-reversed.

Chapter 10

10.2 You get an icosahedron. (An icosahedron is a regular polyhedron with twenty triangular faces.) The icosahedron approximates the geometry of a sphere in exactly the same way hyperbolic paper approximates the geometry of the hyperbolic plane.

10.3 $A = \pi - (\pi/3 + \pi/4 + \pi/6) = \pi - 3\pi/4 = \pi/4.$

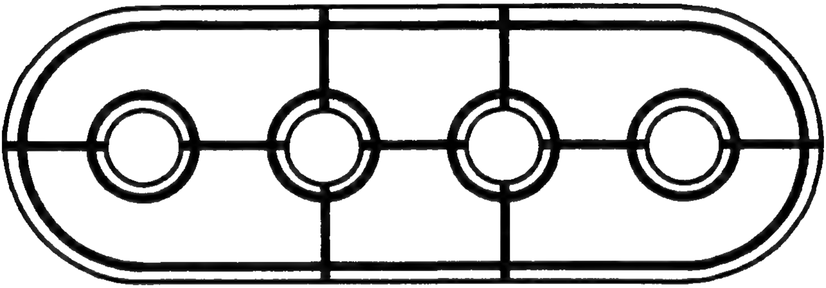
Chapter 11

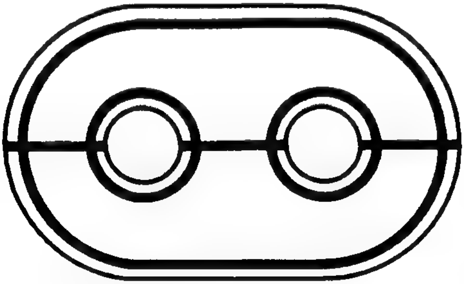
11.1 To eliminate the cone points, put the square on a sphere and let it expand until its corners fit properly (this occurs when the square fills an entire hemisphere and each corner angle is 180°). The surface now has a homogeneous elliptic geometry. You should recognize it as a projective plane (it had the global topology of a projective plane all along, but it just now got the right geometry). The first surface in Figure 11.1 is a projective plane too.

11.2 *First surface*: All six corners come together, hyperbolic geometry eliminates the opposite-of-a-cone-

point. *Second surface*: The corners meet in two groups of three, there are no cone points so the surface has a Euclidean (= flat) geometry. *Third surface*: All eight corners come together, hyperbolic geometry eliminates the opposite-of-a-cone-point. *Fourth surface*: All eight corners come together, hyperbolic geometry eliminates the opposite-of-a-cone-point.

11.3





When you cut an n-holed doughnut surface into hexagons whose corners meet in groups of four you invariably get $4n - 4$ hexagons. (Even though there are different ways to do the cutting, you always get the same number of hexagons!)

11.4 P^2 has an elliptic geometry to begin with. P^2 # P^2 can be given a Euclidean geometry, because the 90° corners of flat squares fit nicely in groups of four. (P^2 # P^2 is topologically the same as K^2, so we knew all along that this surface could be given a Euclidean geometry.) In all the other cases the polygons' corners must be shrunk to fit together in groups of four, so all the other connected sums of projective planes can be given hyperbolic geometry.

11.5

	orientable	nonorientable
elliptic	S^2	P^2
Euclidean	T^2	$P^2\#P^2$ $(= K^2)$
hyperbolic	$T^2\#T^2$ $T^2\#T^2\#T^2$ etc.	$P^2\#P^2\#P^2$ $P^2\#P^2\#P^2\#P^2$ etc.

Chapter 12

12.1 The cell-division has nine vertices, fifteen edges and seven faces.

12.2 An n-gon can be broken down into $n - 2$ triangles. The sum of the angles of the n-gon is the sum

of the angles of all the triangles put together, and the angles of each triangle add up to π.

12.3 The area of the n-gon is the sum of the areas of the $n - 2$ triangles. The area of each triangle is the sum of its angles minus π. Therefore the area of the n-gon is the sum of all the angles minus $n - 2$ times π.

12.4 The method is the same as for a polygon on a sphere, only now the area of each triangle is $\pi - (\alpha + \beta + \gamma)$ instead of $(\alpha + \beta + \gamma) - \pi$. The final formula is $A = (n - 2)\pi -$ (sum of all angles).

12.7 The only change necessary is to use the formula $A = (n - 2)\pi -$ (sum of all angles) instead of $A =$ (sum of all angles) $- (n - 2)\pi$ in Step 2. This is just the negative of what you had before, so the final formula comes out as $A = -2\pi\chi$.

12.8 Using the cell-division of Figure 12.1(b), $\chi(T^2 \# T^2) = 8 - 16 + 6 = -2$. The cell-division of Exercise 11.3 gets to the same answer via a different route: $\chi(T^2 \# T^2) = 6 - 12 + 4 = -2$.

12.9 $A = -2\pi\chi = -2\pi(-2) = 4\pi$.

12.10 If you use the cell-division from Figure 11.3 you get $\chi = 12 - 24 + 8 = -4$. (Other cell-divisions will have different numbers of vertices, edges and faces, but they all give the same answer for the Euler number.) Area $= -2\pi\chi = -2\pi(-4) = 8\pi$.

12.11 $\chi(T^2) = v - e + f = 1 - 2 + 1 = 0$. The Klein bottle's Euler number also works out to be zero no matter what cell-division you use.

12.12 The Euler number decreases by two each time. Presumably the Euler number of $T^2 \# T^2 \# T^2 \# T^2$ is -6.

12.13 The reasoning is the same as that used to compute the Euler number of a connected sum of tori. In the cell-division for a connected sum of projective planes. $v = 2n \times 2 \div 4 = n$, $e = 2 \times 2n \div 2 = 2n$ and of course $f = 2$. So $\chi = v - e + f = n - 2n + 2 = 2 - n$.

Euler number	orientable	nonorientable
2	S^2	
1		P^2
0	T^2	$P^2\#P^2$
-1		$P^2\#P^2\#P^2$
-2	$T^2\#T^2$	$P^2\#P^2\#P^2\#P^2$
-3		$P^2\#P^2\#P^2\#P^2\#P^2$
-4	$T^2\#T^2\#T^2$	etc.
-5		
-6	$T^2\#T^2\#T^2\#T^2$	
-7		
-8	$T^2\#T^2\#T^2\#T^2\#T^2$	
.		
.	etc.	
.		

(Column heading: *orientability*)

12.14 *Figure 11.1*: The first surface has $\chi = 3 - 3 + 1 = 1$, is nonorientable, and is therefore P^2. The second surface has $\chi = 2 - 3 + 1 = 0$, is nonorientable, and is therefore $P^2 \# P^2$ (i.e. it's topologically a Klein bottle). The third surface has $\chi = 1 - 3 + 1 = -1$, is nonorientable, and is therefore $P^2 \# P^2 \# P^2$. *Figure 11.2*: The first surface has $\chi = 1 - 3 + 1 = -1$, is nonorientable, and is therefore $P^2 \# P^2 \# P^2$. The second surface has $\chi = 2 - 3 + 1 = 0$, is nonorientable, and is therefore $P^2 \# P^2$. The third surface has $\chi = 1 - 4 + 1 = -2$, is orientable, and is therefore $T^2 \# T^2$. The fourth surface has $\chi = 1 - 4 + 1 = -2$, is nonorientable, and is therefore $P^2 \# P^2 \# P^2 \# P^2$.

12.15 Curvature is $k = 1/r^2$, so it is measured in (meters)$^{-2}$. The surface 's area A is measured in (meters)^2, so kA is a dimensionless quantity, as is $2\pi\chi$.

12.16 The curvature of a projective plane is the same as the curvature of a sphere of the same radius; in this case $k = 1/r^2 = 1/(2 \text{ meters})^2 = \frac{1}{4}\text{ m}^{-2}$. The area of a projective plane is half the area of a sphere of the same radius; in this case $A = 2\pi r^2 = 2\pi(2 \text{ m})^2 = 8\pi \text{ m}^2$. A projective plane's Euler number is 1 no matter what the radius is, so the Gauss–Bonnet formula reads $(\frac{1}{4}\text{ m}^{-2})(8\pi \text{ m}^2) = 2\pi(1)$.

12.17 Use the formula $kA = 2\pi\chi$ as usual. Plug in $k = -.00001 \text{ m}^{-2}$ and $\chi = -4$, and solve to get $2{,}500{,}000 \text{ m}^2 = 2.5 \text{ km}^2$.

12.18 Plug $k = -3.1658 \times 10^{-6}$ m ^{2} and $A =$ 1,984,707 m^2 into $kA = 2\pi\chi$ and solve for $\chi = -1$. Consult the table in Figure 12.3 to see that the only surface with $\chi = -1$ is P^2 # P^2 # P^2. (It's amazing how far you can get on so little initial information.)

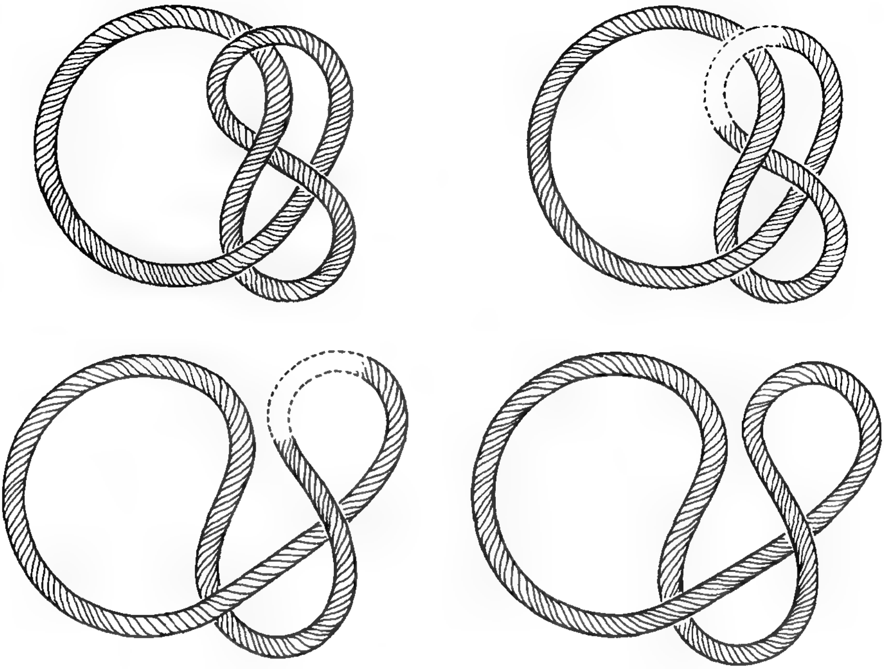
12.19 The T^2 # T^2 is more curved. You can compute the curvature of each surface directly, or if you want to get fancy you can reason that the T^2 # T^2 has twice the Euler number but less than twice the area so it must have a greater curvature.

12.20 Plug into $kA_\Delta = (\alpha + \beta + \gamma) - \pi$ to solve for $k = -1.02 \times 10^{-4}$ km^{-2}. Then use $kA = 2\pi\chi$ to compute $\chi = -4.05$. Presumably χ is -4, and the Flatlanders' (orientable!) universe has the global topology of T^2 # T^2 # T^2 (cf. the table in Figure 12.3). The error in the computed value of χ is most likely due to the uncertainty in the total area.

12.21 The center of the blip has positive curvature (it's convex), but the periphery has negative curvature (it's locally saddle shaped).

Chapter 13

13.1 A spirit could pass one section of rope "over" another section in the fourth dimension. See page 337. The effect is the same as if the one section passed through the other. The knotted loop becomes unknotted.



13.2 Push the "neck" of the Klein bottle into the fourth dimension.

Chapter 14

14.3 You will find yourself trapped on what used to be the outside of the balloon, but is now effectively its inside! What do you have to do to escape? (Assume you can't fit through the mouthpiece and don't want to damage the balloon.)

14.4 First of all, 10 minutes of arc = $\frac{1}{6}$ of a degree = 0.003 radians. The angles of Gauss' triangle must exceed π by at least 0.003 radians if Gauss is to detect the difference with his equipment. The area of the triangle is about 4000 km^2, so Gauss can use the formula

$kA_\Delta = (\alpha + \beta + \gamma) - \pi$ to deduce that the curvature k must be at least $[(\alpha + \beta + \gamma) - \pi]/A_\Delta = [0.003]/(4000 \text{ km}^2) = 0.000001 \text{ km}^{-2}$ if he is to detect it. This corresponds to a universe of radius at most 1000 km! (Use the formula $k = 1/r^2$.) The real universe is obviously not this small!

14.5 The ends of the cylinder are glued to each other with a 180° rotation (not a flip!) so the glued cylinder is topologically a torus. It is orientable because a Flatlander never returns to his starting point mirror-reversed, and it is two-sided because a three-dimensional ant can't get from one side of the surface to the other.

14.6 A disk passing through the center of the ball forms a projective plane in P^3. Any such projective plane is nonorientable and one-sided. (In an orientable three-manifold—like P^3—every orientable surface is two-sided and every nonorientable surface is one-sided.)

Chapter 15

15.1 Our universe might have the local geometry of H^3, but its curvature would be so slight that we haven't yet detected it. (By the way, what is the smallest negative curvature that Gauss could have detected? See Exercise 14.4.)

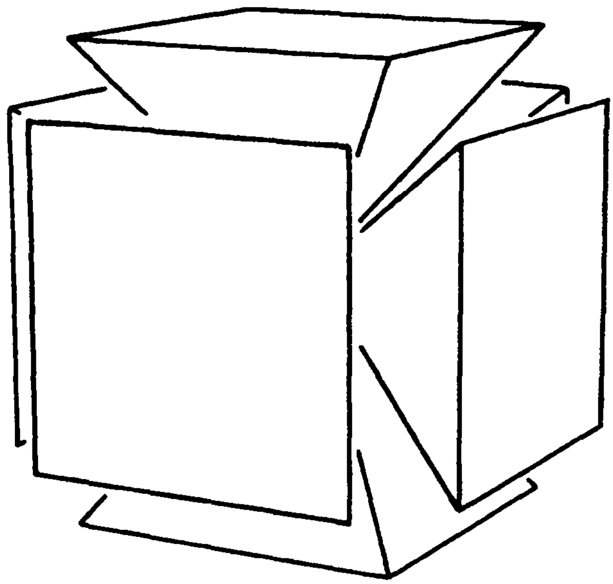
15.2 It would appear infinite to them. From experience their brains would accurately judge distances based on how cross-eyed their eyes are. How would our universe appear to them?

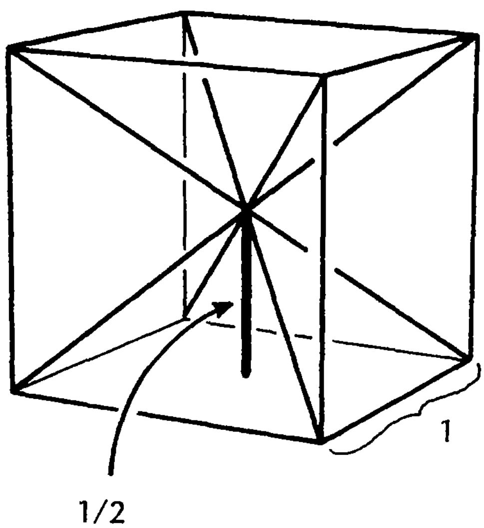
Chapter 16

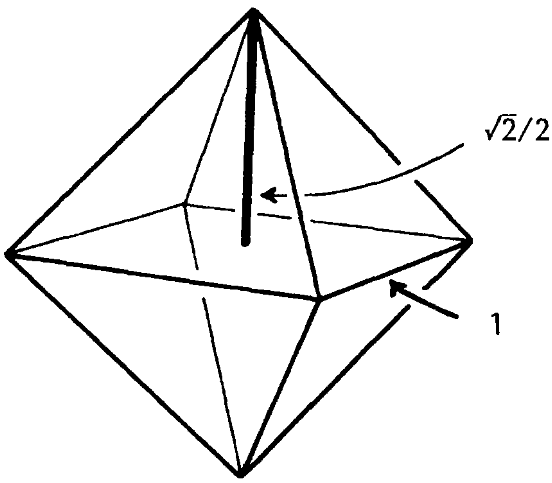
16.1 All four corners come together in one group. (The way you can tell is that when the gluings are taken into account each corner is "adjacent" to the other three.) The corners are too pointy to fit together snugly, so the tetrahedron must be expanded in a hypersphere. This gives the manifold an elliptic geometry.

16.2 The cube's corners come together in two groups of four corners each. (Each corner is adjacent to the other three in its group, but to no others.) The corners are too small to fit snugly in groups of four, and the manifold ends up with an elliptic geometry.

16.3 The octahedron's six corners meet in a single group. The exact pattern is shown in the first two drawings on the next page. Each corner is adjacent to four of the remaining five corners, and each of those four is adjacent to the fifth. The question now is whether the corners are too skinny, too fat, or just right. The corners would be just right if each were exactly as pointy as one of the six pyramids which comprise the cube in the middle drawing below. One way to tell (certainly not the only way) is to use the Pythagorean theorem to discover that the top half of an octahedron is a pyramid whose altitude is about 0.7 times the width of the base, while each of the six pyramids in the middle drawing has an altitude that is exactly 0.5 times the width of its base. This tells you that the octahedron's corners are a little too pointy, so the octahedral space ends up with an elliptic geometry.

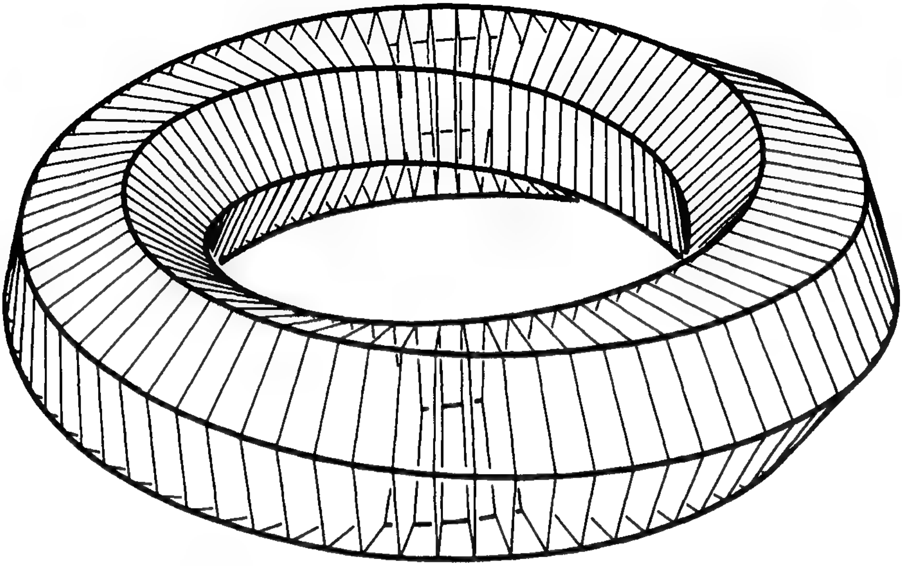


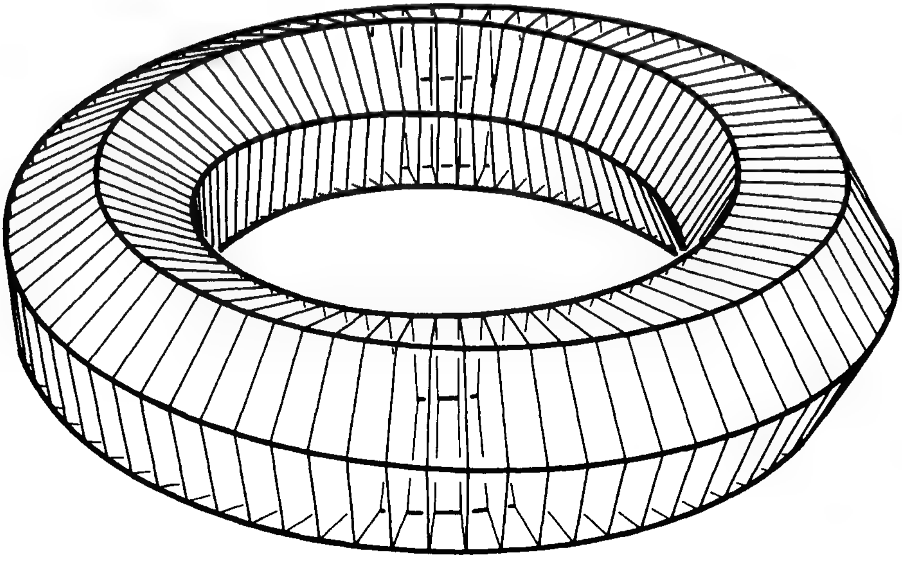

1
1/2


$\sqrt{2}/2$
1

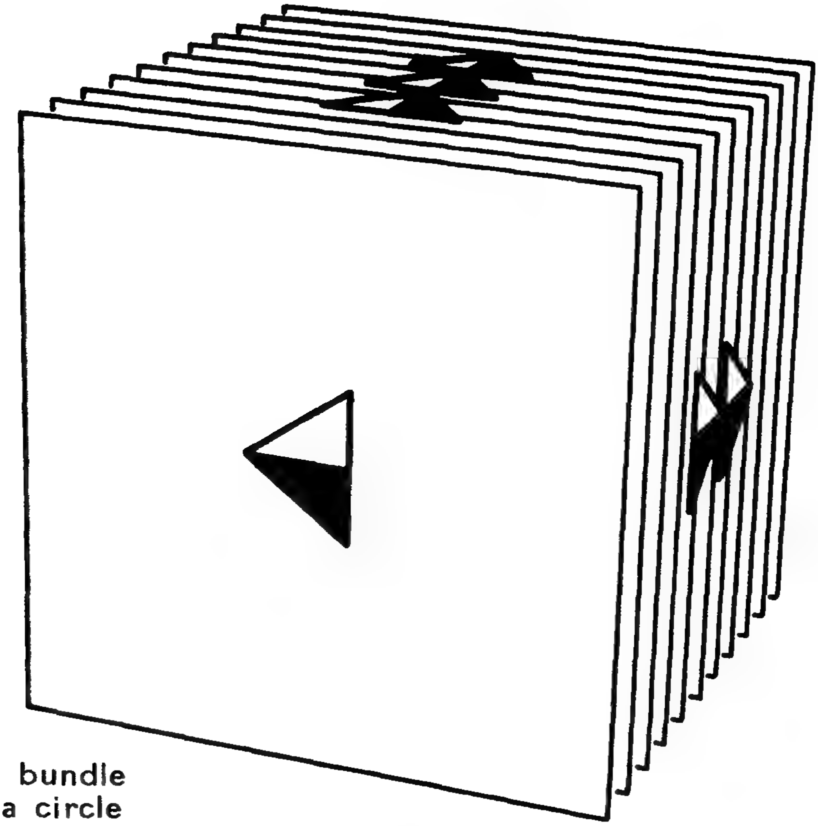
Chapter 17

17.1

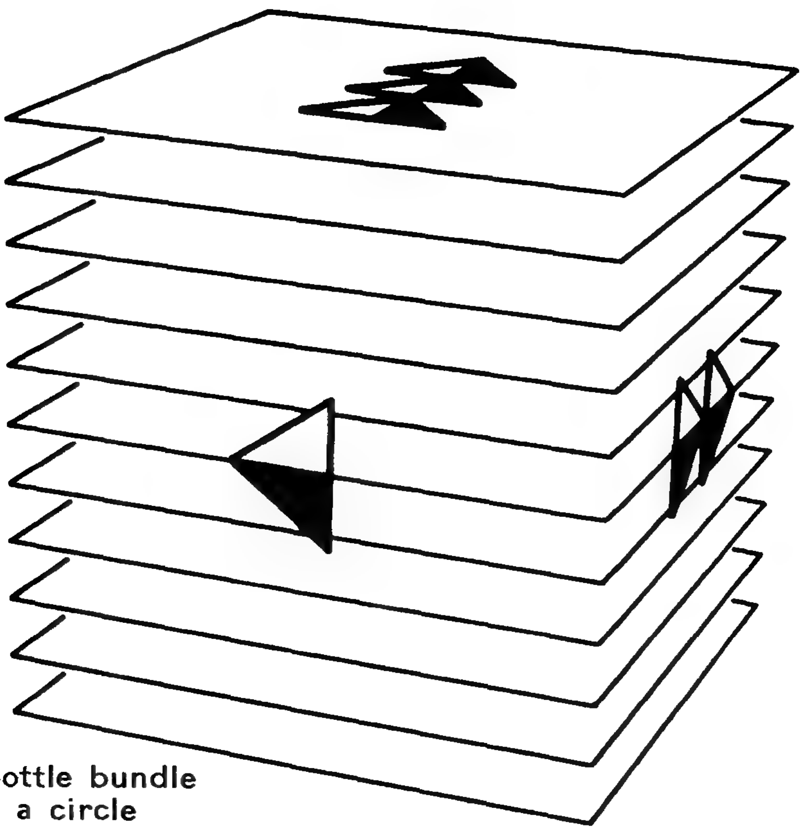




17.2

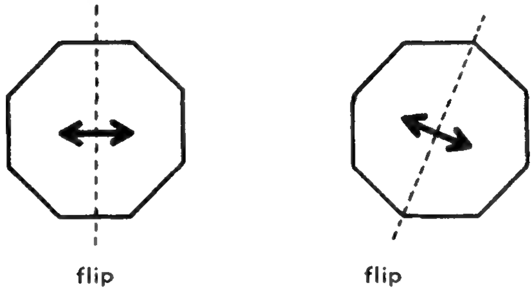


torus bundle
over a circle

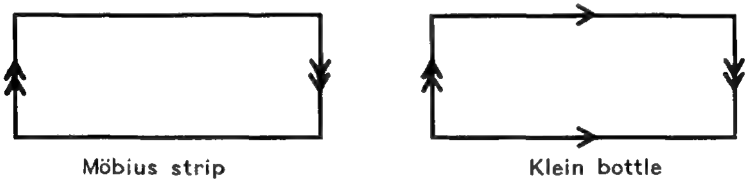


Klein bottle bundle
over a circle

17.3 You can glue the top to the bottom with a $\frac{1}{8}$, $\frac{1}{4}$, $\frac{3}{8}$ or $\frac{1}{2}$ turn. You can also glue them with one of the two types of flips shown below, although for the flips you can't make a nice picture in three-dimensional space like Figure 17.4. (Interestingly enough, the two flips give topologically different manifolds.)



17.4 A torus and a Klein bottle. By the way, a Möbius strip with opposite edges glued is a Klein bottle:



17.5 Start with a solid cylinder and imagine one end to be glued to the other with a side-to-side flip. What you get is a circle of disks, but it's not a product. (You won't be able to physically carry out the gluing in three-dimensional space, just as a Flatlander can't physically assemble a Möbius strip in the plane.)

17.6 Imagine $S^2 \times I$ as a thickened spherical shell. You can glue the inner surface to the outer surface in

the most straightforward way to get $S^2 \times S^1$, or you can glue the inner surface to the outer surface with a side-to-side flip to get a nonorientable sphere bundle over a circle (K^3). Any other way of gluing the inner surface to the outer surface—such as gluing each point on the inner sphere to the diametrically opposite point on the outer sphere—gives a manifold that is topologically the same as one of the two already mentioned (it's like gluing the ends of a cylinder with a fraction of a turn to get a surface that is geometrically different from, but topologically the same as, the torus obtained by gluing the cylinder's ends without the turn; to get a topologically different surface you have to do something more drastic, like glue the cylinder's ends with a side-to-side flip).

You can make a "hypersolid" $S^2 \times S^1$ by gluing together the ends of a solid ball cross an interval in the simplest way. To make a "hypersolid" K^3 you'd have to glue them together with a side-to-side flip. To understand these four-dimensional manifolds, think about how A Square might try to understand an ordinary solid torus or solid Klein bottle.

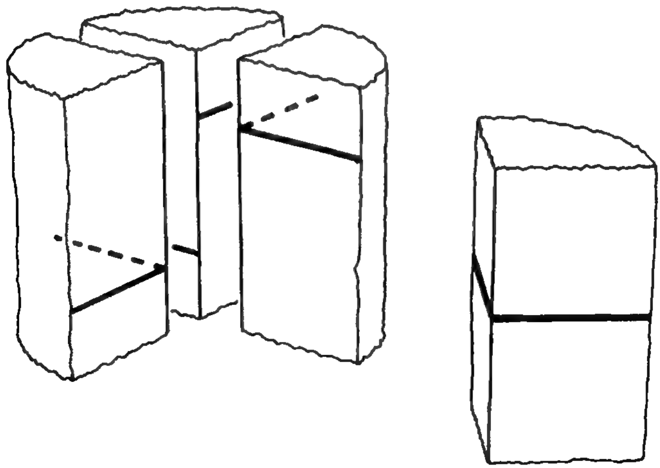
17.7 The manifolds of Exercise 7.4 plus $K^2 \times S^1$. (It turns out that the manifold of Figure 7.10 is topologically equivalent to $K^2 \times S^1$, while the manifolds of Figures 7.11 and 7.12 are topologically equivalent to each other, even though all four are geometrically distinct.)

17.8 Glue the top of the prism to the bottom to get an octagon of circles, and then glue opposite sides to

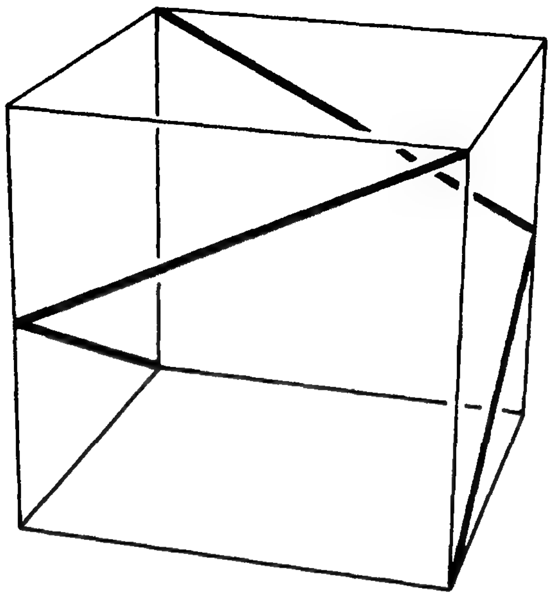
get a two-holed doughnut of circles. (An octagon with opposite sides glued has the topology of $T^2 \# T^2$.)

17.9 Imagine each cube in Figure 8.2 to be filled with vertical spaghetti. When a cube's top is glued to its bottom the spaghetti become circles, so you will get some sort of circle bundle in each case. The first is an orientable circle bundle over a torus, the second is a nonorientable circle bundle over a Klein bottle, the third is a nonorientable circle bundle over a torus, and the fourth is an orientable circle bundle over a Klein bottle.

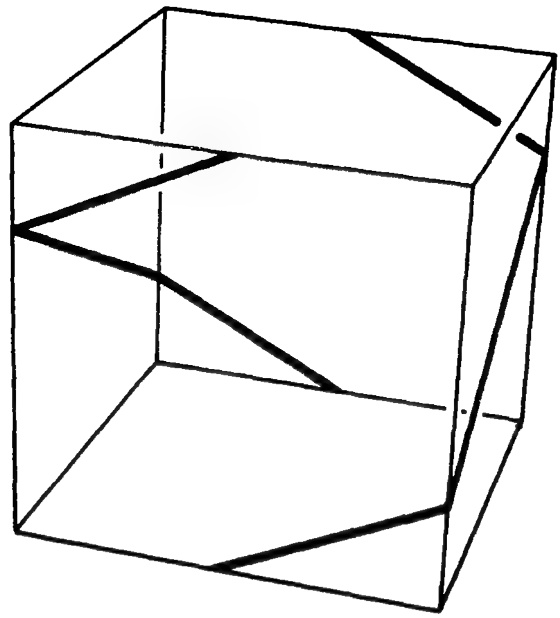
17.10 Any three vertical edges can be glued together so that the horizontal lines match up, but you can never glue in the fourth edge because it has to be at two different "heights" at once. In the regular twisted torus this problem does not arise.



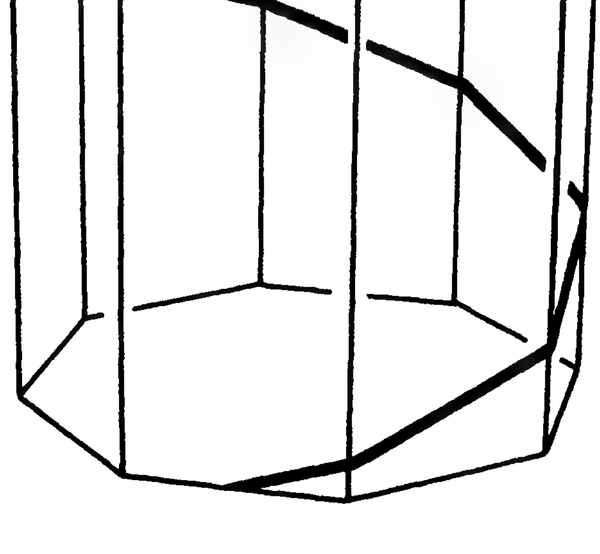
17.11



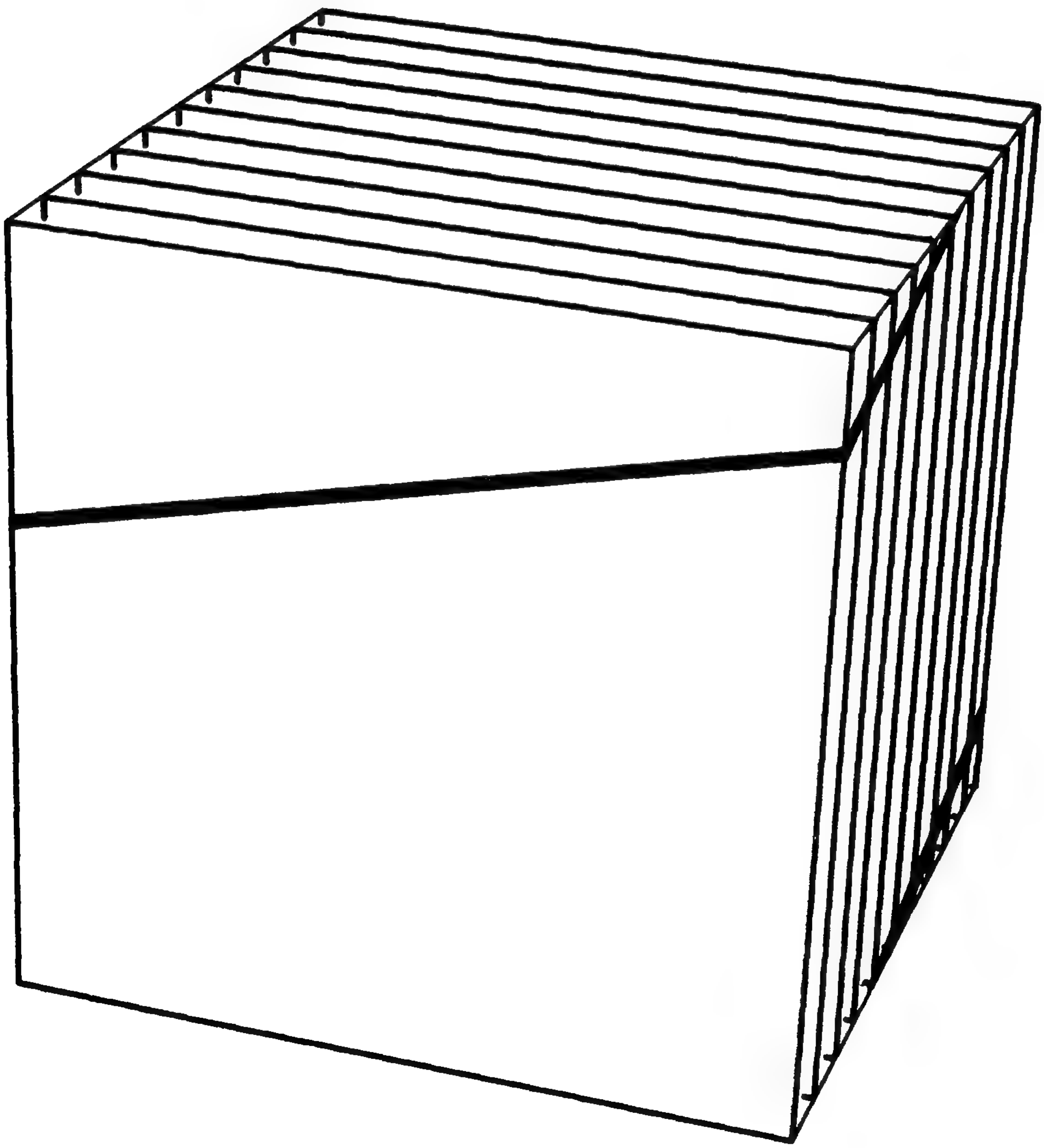
or



17.12



17.13 Surprisingly, the twisted torus *is* a torus bundle over a circle: just slice it up as shown below and note that each slice is topologically a torus in spite of the shear. The twisted torus still isn't a product, though, because the way it's a circle of tori is unrelated to the way it's a torus of circles.



Chapter 18

18.1 K^3 (see Exercise 17.6).

18.2 Any surface cross a circle, just so the surface isn't S^2, P^2, T^2 or K^2. Many surface bundles over circles also have $H^2 \times E$ geometry; see Exercise 17.3 for some specific examples.

18.3 K^3 and $P^2 \times S^1$ are nonorientable, while $S^2 \times S^1$ and the new manifold are orientable. The new manifold is in some ways similar to projective three-space: when you leave either the inner or the outer surface of $S^2 \times I$ (thought of as a thickened spherical shell) you come back rotated, but not mirror reversed, at the diametrically opposite point on the same surface.

18.4 No. In a three-manifold with twisted Euclidean geometry you can easily observe whether you have to travel clockwise or counterclockwise to come back below where you started. Say you are in such a manifold, and say you observe that traveling horizontally in a small clockwise circle brings you back below where you started. A friend of yours sets off on a long journey through distant regions of the manifold. All along the way he continues to observe that motion in a small clockwise circle brings one back below where one starts. When he returns from the long journey he still observes that clockwise motion brings one back below where one starts. This means that he isn't mirror-reversed! For if he were mirror-reversed it would seem to him that motion in a *counterclockwise* circle would

be required to bring one back below where one started. This proves that a traveler never comes home mirror-reversed, so the manifold must be orientable.

18.5 The manifold from Exercise 17.12 will do nicely.

Chapter 19

19.1 The distance between galaxies A and B will increase by (0.07)(15 billion light-years) = 1.05 billion light-years, and so the total separation will be 15 billion light-years + 1.05 billion light-years = 16.05 billion light-years. Galaxy B is moving away from galaxy A at a rate of (1.05 billion light-years)/(1 billion years) = 1.05 light-years per year. This is 5% faster than the speed of light. Fortunately this does not contradict Einstein's theory of special relativity, which says only that two *nearby* objects cannot pass each other at speeds exceeding the speed of light (in technical terms, the relative speed of two objects in the same inertial reference frame cannot exceed the speed of light). In the present example the total distance between the two distant galaxies is increasing at a rate exceeding the speed of light, but that's OK because the two galaxies are very far apart and no direct comparison is possible (in technical terms, the two galaxies lie in different inertial reference frames).

Chapter 21

21.1 The light from a galaxy that is presently a billion light-years away from us has taken somewhat less than a billion years to reach us. The reason is that the

universe was smaller in the past than it was now. Imagine a sequence of equally spaced mileposts: milepost 0 is at the source galaxy, milepost 1 floats in space exactly $1/10$ of the way from the source galaxy to us, milepost 2 floats in space exactly $2/10$ of the way from the source galaxy to us, and so on up to milepost 10 which sits on Earth. In the modern universe the distance between successive mileposts is exactly 100 million light-years. But back when the photon of light was traveling from, say, milepost 8 to milepost 9, the universe was a bit smaller and so the distance between the mileposts was less than 100 million light-years. Further back in the past, when the photon was traveling from milepost 7 to milepost 8, the universe—and the distance between mileposts—was smaller still, and so on. In effect the photon made better progress during the early parts of its journey because space was more compressed back then.

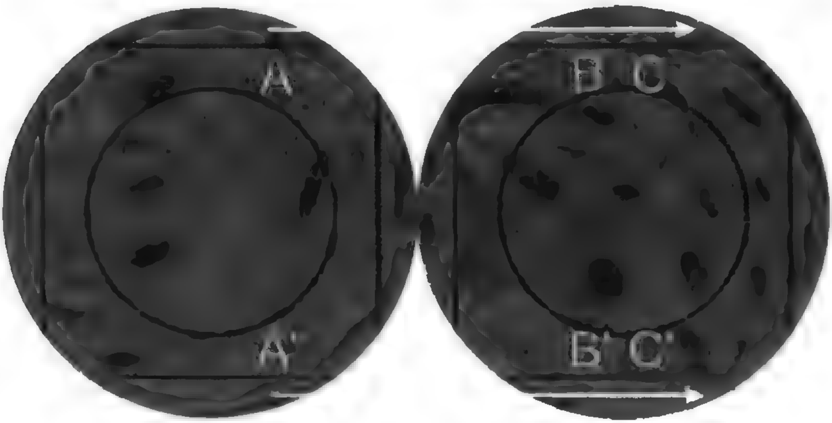
21.2 We may compute the distance from each of the n galaxies to each of the $n - 1$ other galaxies, for a total of $n(n - 1)$ distances. However, the distance from galaxy A to galaxy B is the same as the distance from galaxy B to galaxy A, and so there is no need to compute them separately. This halves the number of distances that must be computed, for a final total of $n(n - 1)/2$ distances.

Chapter 22

22.1 Our last scattering surface will be larger in the future, not only because the universe is expanding but

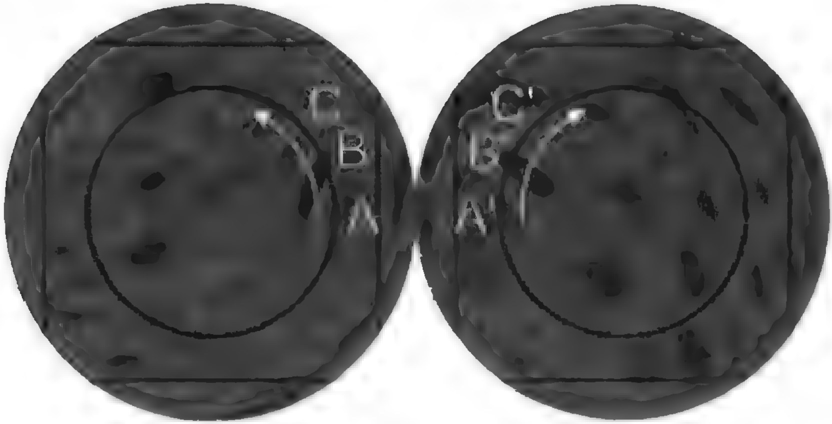
also because light from more distant regions of space will have had time to reach us. In other words, in the future we will be able to see a larger percentage of the universe than we see now. (So if present efforts to detect the topology of space fail, we can wait a billion years and try again!)

22.2 (a) Yes, your fingers pass over identical colors at corresponding points. The temperature at A is the same as the temperature at A′, the temperature at B is the same as at B′, and so on. (b) The temperatures along the circles in the northern and southern skies also match exactly. Note that these two circles, while perfectly round in the sky itself, appear as horizontal lines in the figure below.

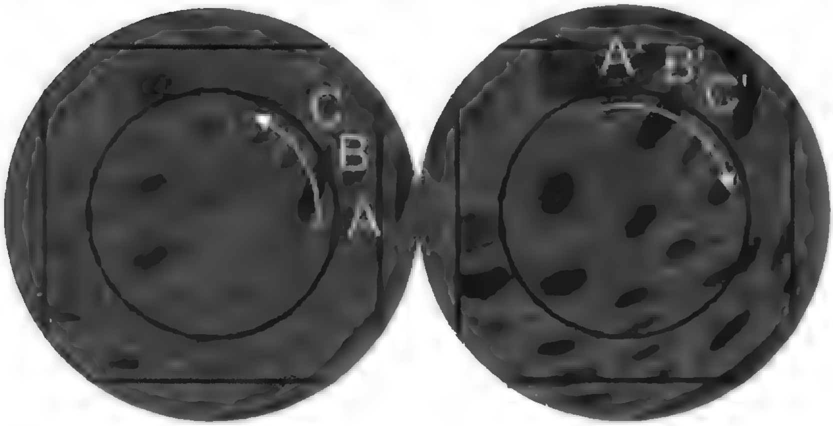


22.3 (a) The circles in a quarter turn manifold match like the circles in a three-torus, except that the three-torus's circles match straight across, while one pair of

the quarter turn manifold's circles match with a 90° rotation. (b) In $K^2 \times S^1$ one pair of circles matches with a flip, thus revealing the topology of the space.



Circles in a three-torus match straight across.



One of the pairs of circles in a quarter-turn manifold matches with a 90° offset.

One of the pairs of circles in $K^2 \times S^1$ matches with a flip. (You'll see the flip most clearly if you imagine bringing the two hemispheres together to restore the spherical sky. Even though the two circles both run counterclockwise on the printed page, on the spherical sky they run in opposite directions.)

22.4 If we're sitting at the center of the dodecahedron, then in both the Poincaré dodecahedral space and the Seifert–Weber space we see six pairs of matching circles (twelve circles total) arranged in the sky with dodecahedral symmetry. The circles nevertheless reveal the difference between the two spaces: in the Poincaré dodecahedral space the temperatures on opposite circles match with a $\frac{1}{10}$ twist, while in the Seifert–Weber space the temperatures match with a $\frac{3}{10}$ twist.

This Page Intentionally Left Blank

Appendix B

Bibliography

FOUR-DIMENSIONAL SPACE

1. Abbott, E.A. Flatland: A Romance of Many Dimensions, Dover, 1992 (originally published 1884).

   If you read only one book from this bibliography, read *Flatland*. Abbott spins together engaging Victorian social satire with a gentle introduction to higher dimensions. The clear and powerful writing is an inspiration in itself. Best of all, this book costs only $1, so when you order your copy

you can get a few for your friends at the same time.

2. Banchoff, T.F. Beyond the Third Dimension: Geometry, Computer Graphics, and Higher Dimensions, Scientific American Library, 1990.
This well illustrated book offers a fun, easy-to-read introduction to four-dimensional space and related ideas in geometry and topology.

3. Rucker, R. Geometry, Relativity, and the Fourth Dimension, Dover, 1977.
An excellent popular exposition of higher dimensions, physics, and philosophy, with a counter-culture twist.

GEOMETRY

4. Hilbert, D. and S. Cohn-Vossen. Geometry and the Imagination, Chelsea, 1999 (translation of Anschauliche Geometrie, 1932).

5. Coxeter, H.S.M. Introduction to Geometry, 2nd edition, Wiley, 1989.
These two books are good general references on geometry. You can read them cover-to-cover, but I prefer to browse around, reading whatever sections catch my interest.

6. Henderson, D.W. Experiencing Geometry, 2nd edition, Prentice Hall, 2001.
An excellent undergraduate geometry textbook, designed for an active learning approach. In-

cludes a comprehensive 26-page bibliography on all aspects of geometry, which the author keeps up to date at *www.math.cornell.edu/~dwh/biblio*.

TOPOLOGY

7. Seifert, H. and W. Threlfall. A Textbook of Topology, Academic Press, 1980 (translation of Lehrbuch der Topologie, 1934).
   Don't be put off by the age of this book. The ideas it contains (homology theory, fundamental group, covering spaces, 3-manifolds, Poincaré duality, linking numbers) are as interesting and relevant today as they were in 1934. An excellent choice for graduate students or advanced undergraduates wanting to get in touch with the origins of 3-manifold topology.

8. Thurston, W.P. Three-Dimensional Geometry and Topology, Vol. 1, Princeton University Press, 1997.
   Bill Thurston pioneered the geometrical theory of 3-manifolds. This book lets you share his vision. The book is intended for graduate students, and you'll need a solid grounding in mathematics to understand all of it. In spite of the stiff prerequisites, Thurston put a lot of energy into making the basic ideas clear, and providing helpful illustrations wherever possible.

9. Weeks, J.R. Exploring the Shape of Space, Key Curriculum Press, 2001.

This 2-week classroom unit uses paper-and-scissors activities, pencil-and-paper games, computer games, and the Shape of Space video to explore possible shapes for our 3-dimensional universe. For students in grades 6–10. *www.keypress.com/space.*

10. Adams, C.C. Knot Book: An Elementary Introduction to the Mathematical Theory of Knots, Freeman, 1994.
A fun survey of knot theory for a broad audience.

11. Rolfsen, D. Knots and Links, Publish or Perish, 1976.
If you're a graduate student in mathematics, *Knots and Links* is an excellent sequel to *The Shape of Space*. Pictures and hands-on exercises take you on a comprehensive yet concrete tour of 3-manifold topology and knot theory. Prerequisites: abstract algebra and some algebraic topology.

12. Francis, G.K. A Topological Picturebook, Springer-Verlag, 1987.
The master of mathematical illustration shows you how to make beautiful and effective drawings, while simultaneously teaching you some classic topology.

13. Guilleman, V. and A. Pollack. Differential Topology, Prentice-Hall, 1974.

This excellent introduction to differential topology explores many beautiful topics—my favorites are vector fields and degrees of mappings—yet it requires only a minimal background, namely calculus and linear algebra (believe me, compared to other books on the topic this is a minimum).

14. Petit, J.-P. Here's Looking at Euclid. Kaufmann, 1985 (translation of Le Geometricon).
This comic book takes a playful romp through some simple 3-manifolds. Novices and experienced mathematicians will all have fun.

COSMIC TOPOLOGY

15. Osserman, R. Poetry of the Universe, Anchor Books, 1996.
This inexpensive paperback tells the story of humanity's evolving understanding of the shape of our universe. Osserman skillfully weaves geometrical understanding into the historical tale. This is truly a layperson's account—no science or mathematics background is needed.

16. Luminet, J.-P. L'Univers Chiffonné, Fayard, 2001.
Clear writing and gorgeous figures make this the best account of current experimental efforts to detect the topology of the real universe. Written for the educated public, the book requires no specialized knowledge of math or physics, and the

style is simple enough that readers with only a minimal command of French will have no trouble. (I still don't understand the word *chiffonné* in the title, but in the text itself the French is easy.) If you can't read French at all, wait for the expected English translation.

17. Cornish, N. and J. Weeks. Measuring the shape of the universe. Notices of the American Mathematical Society 45: 1461–1469, 1998.
Explains the physics and mathematics of an expanding universe, along with the Circles in the Sky method of detecting topology. Written for a broad audience of professional mathematicians, most of the article is accessible to undergraduate math and physics majors as well.

Appendix C

Conway's ZIP Proof

George K. Francis and Jeffrey R. Weeks

Surfaces arise naturally in many different forms, in branches of mathematics ranging from complex analysis to dynamical systems. The Classification Theorem, known since the 1860s, asserts that all closed surfaces, despite their diverse origins and seemingly diverse forms, are topologically equivalent to spheres with some number of handles or crosscaps (to be defined below). The proofs found in most modern text-

This article was originally published in the May 1999 issue of the American Mathematical Monthly.

books follow that of Seifert and Threlfall [5]. Seifert and Threlfall's proof, while satisfyingly constructive, requires that a given surface be brought into a somewhat artificial standard form. Here we present a completely new proof, discovered by John H. Conway in about 1992, which retains the constructive nature of [5] while eliminating the irrelevancies of the standard form. Conway calls it his Zero Irrelevancy Proof, or "ZIP proof," and asks that it always be called by this name, remarking that "otherwise there's a real danger that its origin would be lost, since everyone who hears it immediately regards it as the obvious proof." We trust that Conway's ingenious proof will replace the customary textbook repetition of Seifert–Threlfall in favor of a lighter, fat-free nouvelle cuisine approach that retains all the classical flavor of elementary topology.

We work in the realm of topology, where surfaces may be freely stretched and deformed. For example, a sphere and an ellipsoid are topologically equivalent, because one may be smoothly deformed into the other. But a sphere and a doughnut surface are topologically different, because no such deformation is possible. All of our figures depict deformations of surfaces. For example, the square with two holes in Figure C.1A is topologically equivalent to the square with two tubes (C.1B), because one may be deformed to the other. More generally, two surfaces are considered equivalent, or *homeomorphic*, if and only if one may be mapped onto the other in a continuous, one-to-one

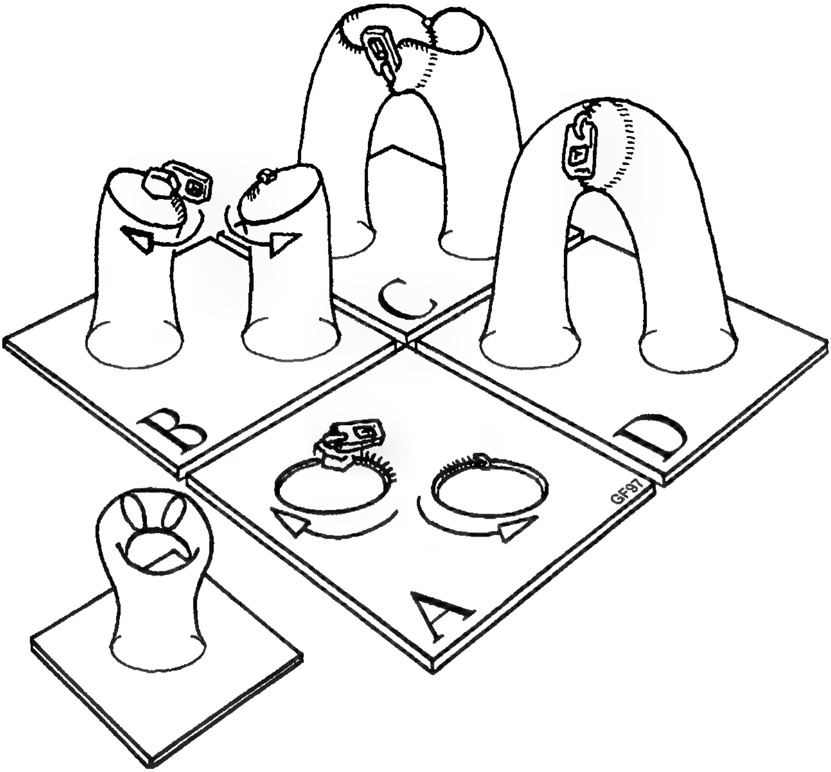


Figure C.1 Handle.

fashion. That is, it's the final equivalence that counts, whether or not it was obtained via a deformation.

Let us introduce the primitive topological features in terms of zippers or "zip-pairs," a zip being half a zipper. Figure C.1A shows a surface with two boundary circles, each with a zip. Zip the zips, and the surface acquires a *handle* (C.1D). If we reverse the direction of one of the zips (C.2A), then one of the tubes must "pass through itself" (C.2B) to get the zip orientations

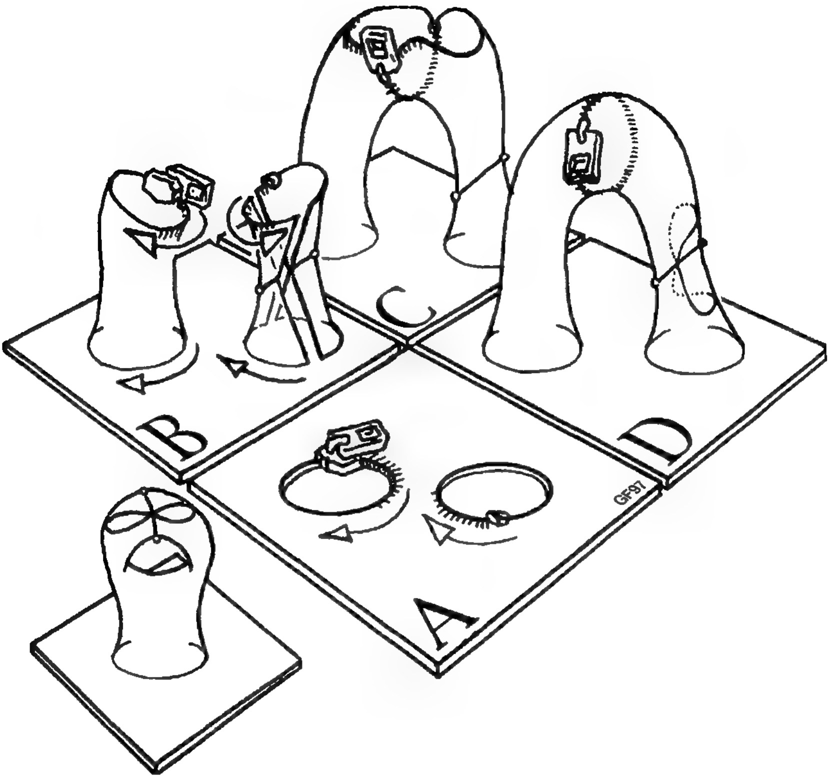


Figure C.2 Crosshandle.

to match. Figure C.2B shows the self-intersecting tube with a vertical slice temporarily removed, so the reader may see its structure more easily. Zipping the zips (C.2C) yields a *crosshandle* (C.2D). This picture of a crosshandle contains a line of self-intersection. The self-intersection is an interesting feature of the surface's placement in 3-dimensional space, but has no effect on the intrinsic topology of the surface itself.

If the zips occupy two halves of a single boundary circle (Figure C.3A), and their orientations are consistent, then we get a *cap* (C.3C), which is topologically trivial (C.3D) and won't be considered further. If the zip orientations are inconsistent (C.4A), the result is more interesting. We deform the surface so that corresponding points on the two zips lie opposite one another (C.4B), and begin zipping. At first the zipper head moves uneventfully upward (C.4C), but upon reaching the top it starts downward, zipping together the "other two sheets" and creating a line of self-intersection. As before, the self-intersection is merely an artifact of the model, and has no effect on the intrinsic topology of the surface. The result is a *crosscap*

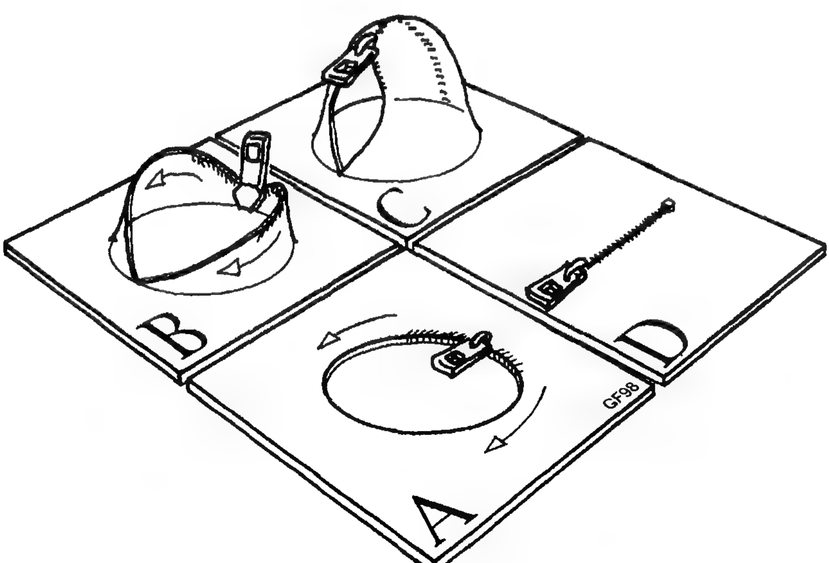


Figure C.3 Cap.

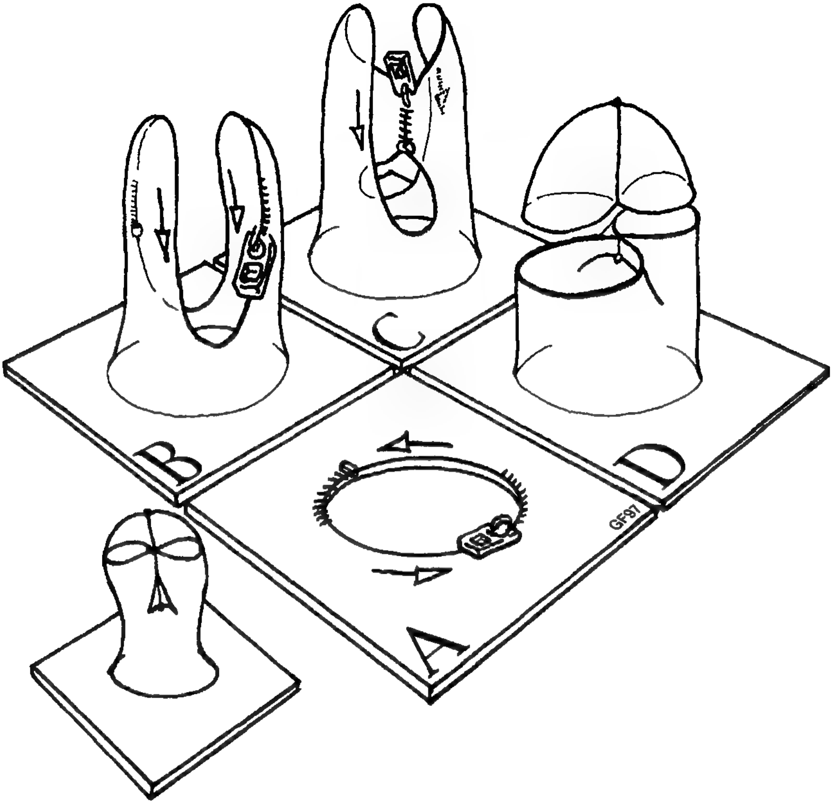


Figure C.4 Crosscap.

(C.4D), shown here with a cutaway view to make the self-intersections clearer.

The preceding constructions should make the concept of a surface clear to non-specialists. Specialists may note that our surfaces are compact, and may have boundary.

Comment. A surface is *not* assumed to be connected.

Comment. Figure C.5 shows an example of a triangulated surface. All surfaces may be triangulated, but the proof [4] is difficult. Instead we may consider the Classification Theorem to be a statement about surfaces that have already been triangulated.

Definition. A *perforation* is what's left when you remove an open disk from a surface. For example, Figure C.1A shows a portion of a surface with two perforations.

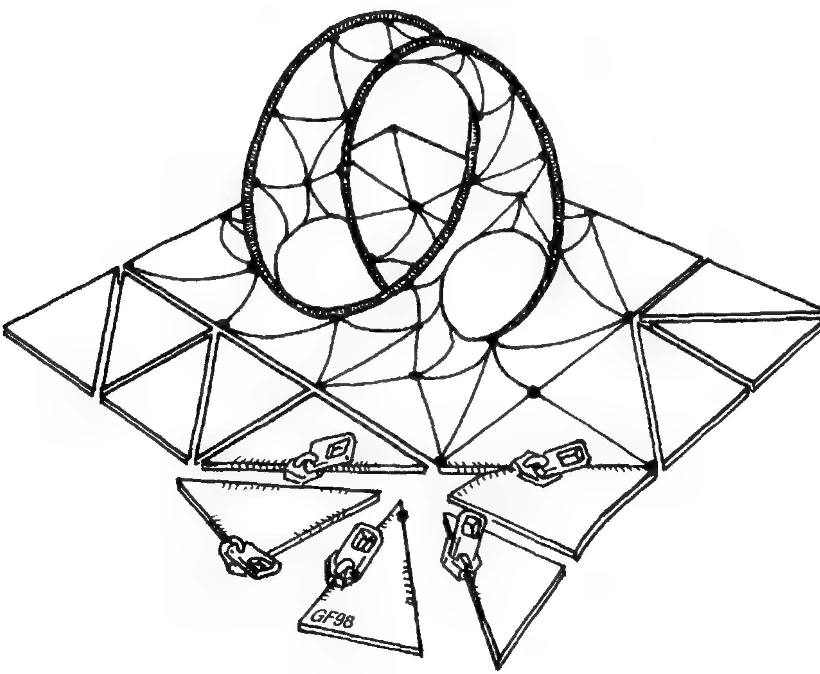


Figure C.5 Install a zip-pair along each edge of the triangulation, unzip them all, and then rezip them one at a time.

Definition. A surface is *ordinary* if it is homeomorphic to a finite collection of spheres, each with a finite number of handles, crosshandles, crosscaps, and perforations.

Classification Theorem (preliminary version)
Every surface is ordinary.

Proof Begin with an arbitrary triangulated surface. Imagine it as a patchwork quilt, only instead of imagining traditional square patches of material held together with stitching, imagine triangular patches held together with zip-pairs (Figure C.5). Unzip all the zip-pairs, and the surface falls into a collection of triangles with zips along their edges. This collection of triangles is an ordinary surface, because each triangle is homeomorphic to a sphere with a single perforation. Now rezip one zip to its original mate. It's not hard to show that the resulting surface must again be ordinary, but for clarity we postpone the details to Lemma 1. Continue rezipping the zips to their original mates, one pair at a time, noting that at each step Lemma 1 ensures that the surface remains ordinary. When the last zip-pair is zipped, the original surface is restored, and is seen to be ordinary. □

Lemma 1. *Consider a surface with two zips attached to portions of its boundary. If the surface is ordinary before the zips are zipped together, it is ordinary afterwards as well.*

Proof First consider the case that each of the two zips completely occupies a boundary circle. If the two boundary circles lie on the same connected component of the surface, then the surface may be deformed so that the boundary circles are adjacent to one another, and zipping them together converts them into either a handle (Figure C.1) or a crosshandle (Figure C.2), according to their relative orientation. If the two boundary circles lie on different connected components, then zipping them together joins the two components into one.

Next consider the case that the two zips share a single boundary circle, which they occupy completely. Zipping them together creates either a cap (Figure C.3) or a crosscap (Figure C.4), according to their relative orientation.

Finally, consider the various cases in which the zips needn't completely occupy their boundary circle(s), but may leave gaps. For example, zipping together the zips in Figure C.6A converts two perforations into a handle with a perforation on top (C.6B). The perforation may then be slid free of the handle (C.6C,6D). Returning to the general case of two zips that needn't completely occupy their boundary circle(s), imagine that those two zips retain their normal size, while all other zips shrink to a size so small that we can't see them with our eyeglasses off. This reduces us (with our eyeglasses still off!) to the case of two zips that *do* completely occupy their boundary circle(s), so we zip them and obtain a handle, crosshan-

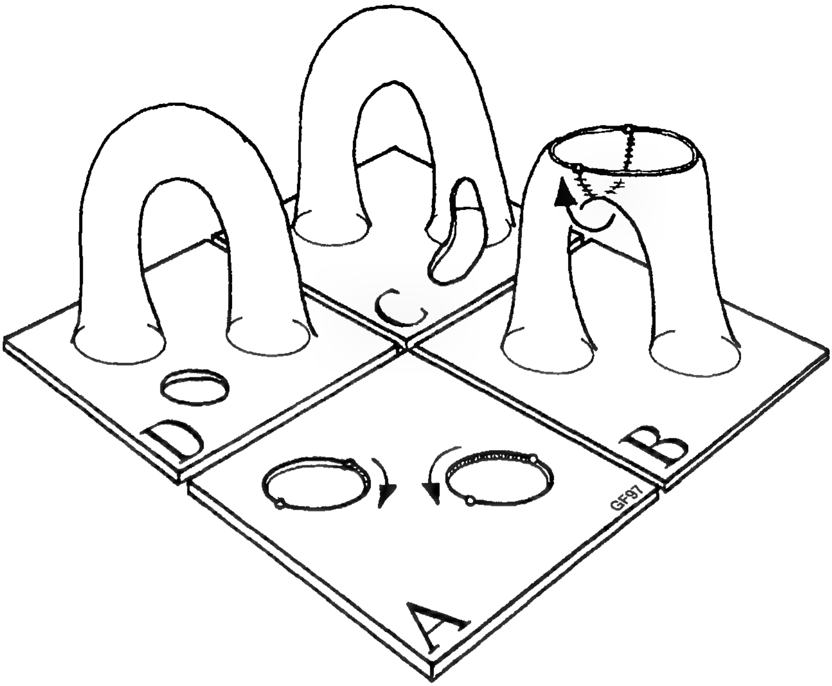


Figure C.6 These zips only partially occupy the boundary circles, so zipping them yields a handle with a puncture.

dle, cap, or crosscap, as illustrated in Figures C.1–4. When we put our eyeglasses back on, we notice that the surface has small perforations as well, which we now restore to their original size. □

The following two lemmas express the relationships among handles, crosshandles, and crosscaps.

Lemma 2. *A crosshandle is homeomorphic to two crosscaps.*

Proof Consider a surface with a "Klein perforation" (Figure C.7A). If the parallel zips (shown with black arrows in C.7A) are zipped first, the perforation splits in two (C.7B). Zipping the remaining zips yields a crosshandle (C.7C).

If, on the other hand, the antiparallel zips (shown with white arrows in Figure C.7A) are zipped first, we get a perforation with a "Möbius bridge" (C.7D). Raising its boundary to a constant height, while letting the surface droop below it, yields the bottom half of a crosscap (C.7E). Temporarily fill in the top half of the crosscap with an "invisible disk" (C.7F), slide the disk

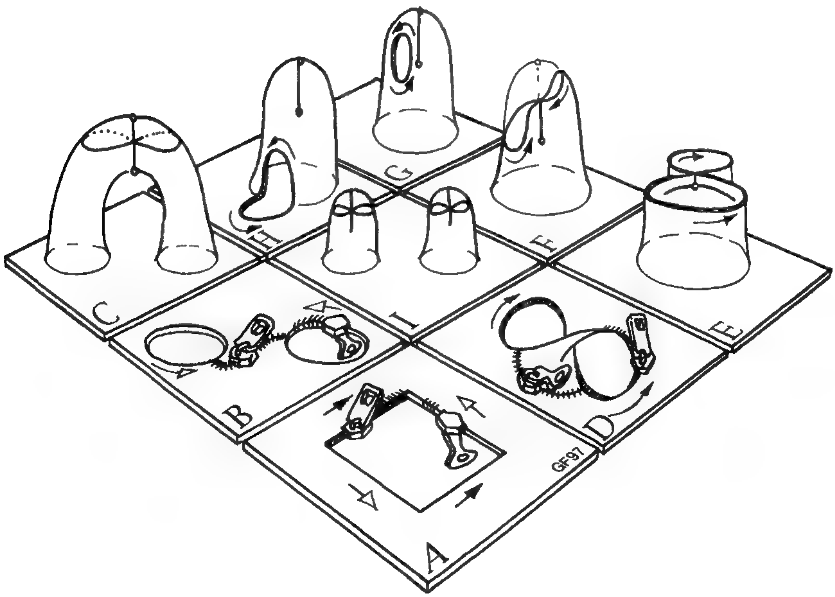


Figure C.7 A crosshandle is homeomorphic to two crosscaps.

free of the crosscap's line of self-intersection (C.7G), and then remove the temporary disk. Slide the perforation off the crosscap (C.7H) and zip the remaining zip-pair (shown with black arrows) to create a second crosscap (C.7I).

The intrinsic topology of the surface does not depend on which zip-pair is zipped first, so we conclude that the crosshandle (C.7C) is homeomorphic to two crosscaps (C.7I). □

Lemma 3 (Dyck's Theorem [1]). *Handles and crosshandles are equivalent in the presence of a crosscap.*

Proof Consider a pair of perforations with zips installed as in Figure C.8A. If, on the one hand, the black arrows are zipped first (C.8B), we get a handle along with instructions for a crosscap. If, on the other hand, one tube crosses through itself (C.8C, recall also Figure C.2B) and the white arrows are zipped first, we get a crosshandle with instructions for a crosscap (C.8D). In both cases, of course, the crosscap may be slid free of the handle or crosshandle, just as the perforation was slid free of the handle in Figure C.6BCD. Thus a handle-with-crosscap is homeomorphic to a crosshandle-with-crosscap. □

Classification Theorem

Every connected closed surface is homeomorphic to either a sphere with crosscaps or a sphere with handles.

Proof By the preliminary version of the Classification Theorem, a connected closed surface is homeo-

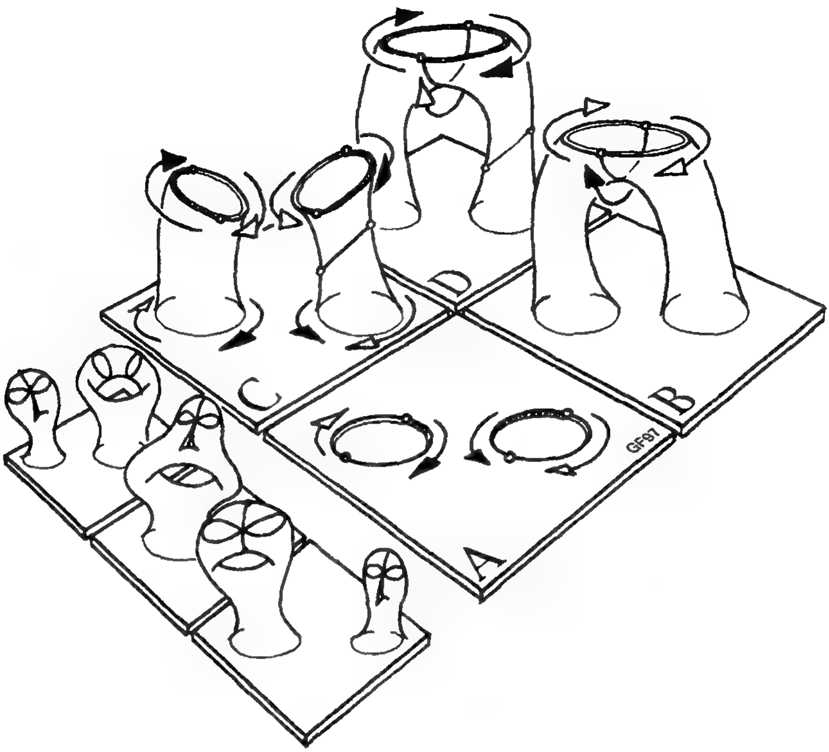


Figure C.8 The presence of a crosscap makes a handle cross.

morphic to a sphere with handles, crosshandles, and crosscaps.

Case 1 At least one crosshandle or crosscap is present. Each crosshandle is homeomorphic to two crosscaps (Lemma 2), so the surface as a whole is homeomorphic to a sphere with crosscaps and handles only. At least one crosscap is present, so each handle is equivalent to a crosshandle (Lemma 3), which is in

turn homeomorphic to two crosscaps (Lemma 2), resulting in a sphere with crosscaps only.

Case 2 No crosshandle or crosscap is present. The surface is homeomorphic to a sphere with handles only.

We have shown that every connected closed surface is homeomorphic to either a sphere with crosscaps or a sphere with handles. □

Comment. The surfaces named in the Classification Theorem are all topologically distinct, and may be recognized by their orientability and Euler number. A sphere with n handles is orientable with Euler number $2 - 2n$, while a sphere with n crosscaps is nonorientable with Euler number $2 - n$. Most topology books provide details; elementary introductions appear in [6] and [2].

Nomenclature. A sphere with one handle is a *torus*, a sphere with two handles is a *double torus*, with three handles a *triple torus*, and so on. A sphere with one crosscap has traditionally been called a real projective plane. That name is appropriate in the study of projective geometry, when an affine structure is present, but is inappropriate for a purely topological object. Instead, Conway proposes that a sphere with one crosscap be called a *cross surface*. The name cross surface evokes not only the crosscap, but also the surface's elegant alternative construction as a sphere

with antipodal points identified. A sphere with two crosscaps then becomes a *double cross surface*, with three crosscaps a *triple cross surface*, and so on. As special cases, the double cross surface is often called a *Klein bottle*, and the triple cross surface *Dyck's surface* [3].

REFERENCES

1. W. Dyck. Beiträge zur Analysis Situs I. Math. Ann. 32: 457–512, 1888.

2. D. Farmer and T. Stanford. Knots and Surfaces, American Mathematical Society, Providence RI, 1996.

3. G. Francis and B. Collins. On knot-spanning surfaces: An illustrated essay on topological art. In: Michele Emmer, ed., The Visual Mind: Art and Mathematics, Chapter 11, MIT Press, Cambridge MA, 1993.

4. T. Radó. Über den Begriff der Riemannschen Fläche. Acta Litt. Sci. Szeged 2: 101–121, 1925.

5. H. Seifert and W. Threlfall. Lehrbuch der Topologie, Teubner, Leipzig, 1934. Translated into English as A Textbook of Topology, Academic Press, New York, 1980.

6. J. Weeks. The Shape of Space, Marcel Dekker, New York, 1985.

This Page Intentionally Left Blank

Index